

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 1 7, 2006, 21:24:06 ; Search time 4213 Seconds
(without alignments)
10038.340 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744

Sequence: 1 gggcgacatcagaagttt.....aaagcagcgtatgctcttcg 744

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hfg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	6	CS130878
2	40.8	5.5	110000	1	BA000004_31
3	40.2	5.4	239687	14	CR932017
4	39.8	5.3	200322	14	AC164371
5	39.4	5.3	2000	6	AX655393
6	39.4	5.3	204839	14	AC144482
7	38.6	5.2	110000	2	AC116984_2
8	38.6	5.2	110000	2	AC116984_3
9	38.6	5.2	163151	5	BX537280
10	38.6	5.2	244259	14	AC123226
11	38.2	5.2	257204	14	AC109987
12	38.2	5.1	116374	5	AC123226
13	38.2	5.1	242081	14	BX537280
14	37.8	5.1	576	15	AF225410
15	37.8	5.1	1554	15	AY032588
16	37.8	5.1	173237	14	AC155465
17	37.8	5.1	181924	14	AC155611
18	37.8	5.1	200412	14	AC155666

c	19	37.8	5.1	279242	14	AC114079	AC114079	Rattus no
	20	37.8	5.1	346208	14	AC128290	AC128290	Rattus no
	21	37.4	5.0	168293	14	AC018397	AC018397	Homo sapi
	22	37	5.0	33961	8	AC000044	AC000044	Homo sapi
	23	37	5.0	43738	8	AC000034	AC000034	Homo sapi
	24	37	5.0	226007	14	AC103426	AC103426	Rattus no
	25	36.6	4.9	2000	6	AX655393	AX655393	Sequence
	26	36.6	4.9	2286	15	LEU78526	U78526	Lycopersicon
	27	36.4	4.9	176422	9	AC158170	AC158170	Mus muscu
	28	36.2	4.9	387	15	AF123309	AF123309	Gasterrodia
	29	36.2	4.9	516	15	GBL277784	AJ277784	Gasterrodia
	30	36.2	4.9	687	15	GBL277785	AJ277785	Gasterrodia
	31	36.2	4.9	687	15	GBL277786	AJ277786	Gasterrodia
	32	36.2	4.9	699	15	GBL277783	AJ277783	Gasterrodia
	33	36.2	4.9	2306	15	AF334813	AF334813	Gasterrodia
	34	36.2	4.9	41741	8	AL356500	AL356500	Human DNA
	35	36.2	4.9	110000	15	AP008217_237	AC109832	Oryza sat
	36	36.2	4.9	149697	15	AC109832	AC109832	Oryza sat
	37	36.2	4.9	180695	14	AC140661	AC140661	Pan trogl
	38	36	4.8	63331	14	AC084345	AC084345	Homo sapi
	39	36	4.8	216589	14	AC091741	AC091741	Homo sapi
	40	35.6	4.8	26557	8	DQ001128	DQ001128	Homo sapi
	41	35.6	4.8	79438	8	HSDJ547C9	AL109927	Human DNA
	42	35.6	4.8	134953	14	AC026088	AC026088	Homo sapi
	43	35.6	4.8	154350	8	AC016948	AC016948	Homo sapi
	44	35.6	4.8	194418	5	BX001055	BX001055	Zebrafish
	45	35.6	4.8	197196	14	AC112585	AC112585	Rattus no

ALIGNMENTS

RESULT 1	CS130878	744 bp	DNA	Linear	PAT 02-AUG-2005
LOCUS	CS130878	Sequence 18 from Patent WO2005066339.			
DEFINITION	CS130878				
ACCESSION	CS130878.1	GI:71793146			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
misc_feature					
misc_feature					
misc_feature					
ORIGIN					
Query Match					
Best Local Similarity					
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
1	GGGCAATTATGAAAGTTTGGGATGATTTGGTACCGTATATGATGCTGTTATTCAG	60			
1	GGGCAATTATGAAAGTTTGGGATGATTTGGTACCGTATATGATGCTGTTATTCAG	60			
61	GGGCGCAATGACAGATTCCTGGCTTTGCCGCTTCAATGTTGGCATTAAATGCTTC	120			
61	TGGGGCGAATGACAGATTCCTGGCTTTGCCGCTTCAATGTTGGCATTAAATGCTTC	120			

Consensus quality: 235832 bases at least Q30
Consensus quality: 236380 bases at least Q20
Insert size: 239087; sum-of-contigs
Insert size: 217257; 9.8% error; agarose-fp
Quality coverage: 9.53x in Q20 bases; sum-of-contigs Quality
coverage: 10.66x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 5067: contig of 5067 bp in length
* 5068 5167: gap of 100 bp
* 5168 41822: contig of 36655 bp in length
* 41823 41922: gap of 100 bp
* 41923 131937: contig of 90015 bp in length
* 131938 132037: gap of 100 bp
* 132038 143502: contig of 11465 bp in length
* 143503 208913: gap of 100 bp
* 208914 209013: contig of 65311 bp in length
* 209014 236615: gap of 100 bp
* 236616 236715: gap of 100 bp
* 236716 239687: contig of 2972 bp in length.

FEATURES

source

1. 239687
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEY-33P9"
/clone_id="DankKey"
1. 5067
/note="assembly fragment:00012
fragment chain:1"
misc_feature
5168. 41822
/note="assembly fragment:00593
fragment chain:1"
misc_feature
41923. 131937
/note="assembly fragment:02166
fragment chain:1"
misc_feature
132038. 143502
/note="assembly fragment:00090
fragment chain:1"
misc_feature
143603. 208913
/note="assembly fragment:01089
fragment chain:2"
misc_feature
209014. 236615
/note="assembly fragment:00175
fragment chain:2"
misc_feature
236716. 239687
/note="assembly fragment:00050.0"

ORIGIN

Query Match 5.4%; Score 40.2; DB 14; Length 239687;
Best Local Similarity 56.4%; Pred. No. 1.2;
Matches 75; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 508 AATATTCATTCGACGAACCTTAACGGAACGTTTATCAGCCGATTTCTACCTGAGATA 567
DB 152572 AATATTCGAATTAAGTTCTTACACGAAATTTACTTAACCAAAATTAAGTTTGAATATA 152631
QY 568 TATGCAATATACAAATTTACAGGACTTATTCGACGACCTCAACCCCTGTGTCTCAATGAGC 627
DB 152632 TTTTCGTTTGAATTTACAGATATTTTTCATGAACTCAACCTTCTTATTAATTTTAC 152691
QY 628 ACAAAATCTCGTAT 640
DB 152692 TGATTTCTGCGCAT 152704

RESULT 4

AC164371

LOCUS

DEFINITION

AC164371 200322 bp DNA linear HTG 01-JUL-2005
Bos taurus clone CH240-11714, *** SEQUENCING IN PROGRESS ***, 19
unordered pieces.

ACCESSION

AC164371.2

GI:68299997

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

KEYWORDS

Bos taurus

SOURCE

Bos taurus

ORGANISM

Bos taurus

REFERENCE

1 (bases 1 to 200322)

Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, B., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Burch, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cre, A., D'Souza, L., Davila, M.L., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregoris, B., Geer, K., Gill, R., Grady, M., Gueter, W., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, K., Hamilton, K., Harevey, Y., Havlak, P., Hawes, S., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladum, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewala, L., Louised, H., Lozano, R.J., Lu, X., Ma, D., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Marshney, S., McLeod, M.P., McKell, T.Z., Meenen, B., Miosavljivic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okwomou, G., Olampunagoun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindecker, A., Popovic, D., Primis, B., Pu, L., Pu, L., Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, B., Song, X.-Z., Sorelle, K., Soza, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabot, J., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 200322)
Worley, K.C.
Direct Submission
Submitted (20-JUN-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 200322)

AUTHORS
TITLE
JOURNAL

COMMENT

Cow Genome Sequencing Consortium.

Direct Submission

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jun 29, 2005 this sequence version replaced gi:67972940.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FHS

Center clone name: CH240-11714

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 190643 bases at least Q40

Consensus quality: 193630 bases at least Q30

Consensus quality: 195478 bases at least Q20

Estimated insert size: 198954; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 19 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence.

* as soon as it is available and the accession number will

* be preserved.

```
1 5877: contig of 5877 bp in length
* 5878 6322: gap of 445 bp
* 6323 18172: contig of 11850 bp in length
* 18173 18222: gap of 50 bp
* 18223 20526: contig of 2304 bp in length
* 20527 20576: gap of 50 bp
* 20577 35742: contig of 15166 bp in length
* 35743 36222: gap of 480 bp
* 36223 38281: contig of 2059 bp in length
* 38282 38381: gap of unknown length
* 38382 41147: contig of 2766 bp in length
* 41148 41197: gap of 50 bp
* 41198 55437: contig of 14240 bp in length
* 55438 55562: gap of 125 bp
* 55563 86280: contig of 30718 bp in length
* 86281 86330: gap of 50 bp
* 86331 92028: contig of 5698 bp in length
* 92029 92078: gap of 50 bp
* 92079 119970: contig of 27892 bp in length
* 119971 120047: gap of 77 bp
* 120048 123375: contig of 3328 bp in length
* 123376 123665: gap of 290 bp
* 123666 142980: contig of 19315 bp in length
* 142981 143080: gap of unknown length
* 143081 144260: contig of 1180 bp in length
* 144261 145079: gap of 819 bp
* 145080 192359: contig of 47280 bp in length
* 192360 192459: gap of unknown length
* 192460 193549: contig of 1090 bp in length
```

FEATURES
source

```
* 193550 193649: gap of unknown length
* 193650 195025: contig of 1376 bp in length
* 195026 195125: gap of unknown length
* 195126 196204: contig of 1079 bp in length
* 196205 196304: gap of unknown length
* 196305 197805: contig of 1501 bp in length
* 197806 197905: gap of unknown length
* 197906 200322: contig of 2417 bp in length.
```

location/Qualifiers

1..200322

/organism="Bos taurus"

/mol_type="genomic DNA"

/db_xref="taxon:9913"

/clone="CH240-11714"

5878..6322

/estimated_length=445

18173..18222

/estimated_length=50

20527..20576

/estimated_length=50

35743..36222

/estimated_length=480

38282..38381

/estimated_length=unknown

41148..41197

/estimated_length=50

55438..55562

/estimated_length=125

86281..86330

/estimated_length=50

92029..92078

/estimated_length=50

119971..120047

/estimated_length=77

123376..123665

/estimated_length=290

142981..143080

/estimated_length=unknown

144261..145079

/estimated_length=819

192360..192459

/estimated_length=unknown

193550..193649

/estimated_length=unknown

195026..195125

/estimated_length=unknown

196205..196304

/estimated_length=unknown

197806..197905

/estimated_length=unknown

ORIGIN

Query Match 5.3%; Score 39.8; DB 14; Length 200322;

Best Local Similarity 60.7%; Pred. No. 1.5;

Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
OY 322 AAGTACGATATATCCGCTTTCACACATACGGAAGTAGAGACTTGAATTAGATT 381
Db 13752 AAGTAGACATACGACAGCTCTTCTCCGGAAGTAGAGACTTGAATTAGAGGA 13811
OY 382 AACTCCCTGACAGCAAGTTAAGAGTGTGGGTTACGACGGAATAGTGA 428
Db 13812 AAAACCTCTCGATATGATACAAATCTGGGCTCTCGGGAAGGGA 13858
```

RESULT 5

AX655393/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

```
AX655393 2000 bp DNA linear PAT 22-MAR-2003
Sequence 5263 from Patent WO03000898.
AX655393
AX655393.1 GI:29158207
Oryza sativa
```


ORGANISM	Oryza sativa
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
AUTHORS	1 Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M., Katagiri,F., Quan,S., Tao,Y., Whitman,S., Xie,Z., Zhu,T. and Zou,G.
TITLE	Plant genes involved in defense against pathogens
JOURNAL	Patent:WO 03000898-A 5263 (3)-JAN-2003; Syngenta Participations AG (CH)
FEATURES	Participations AG (CH) location/Qualifiers
source	1..2000 /organism="Oryza sativa" /mol_type="unassigned DNA" /db_xref="taxon:4530"
ORIGIN	
Query Match	5.3%; Score 39.4; DB 6; Length 2000;
Best Local Similarity	9.1%; Pred. 1.7;
Matches	61; Conservative 304; Mismatches 300; Indels 4; Gaps 2;
Oy	24 ATGATTTGGTACCGGTATATCATCTGTTATTTCAGTGGGGCGAATGACAGAAATCT 83
Db	1002 MMKTKTKMTAVSSWKKYVAWKRAVAMSSSKTKMTCTGKRAATYGTMMAGRRRW 943
Oy	84 GCGTTTCCGCGTTCATTTGGCGATTAATGTCTTCGCGAACACGATATGATTCG 143
Db	942 RMAMCMYCCMMWMTCTSCMMWKYRTWSCVYTMWGMARVAYVAMRRBRTYKMSWRM 883
Oy	144 AGGAAACACACCGCGGAACGCATCAACGCTAACGCTCAACTTTTGGTTAATTCAC 203
Db	882 YMTMTKMAWMTWTCMCARWYMAATGMAWMMWRMYMTYCYAMTCAKCKYMAATMKWTT 823
Oy	204 GAATAGTTTACAGGTAGGACGAAGACGAACAACTCGATAGTCGTCGCGCTAATGTC 263
Db	822 WACAMRATSMWRAMAGRRKRYKAKRYVWRWRCTAGARWKSRYR--WKMKTA 766
Oy	264 TAGCAACTCATTTGCAACATCGGACATCAAGCTGCTACGACTTCATCTTCAGTCGA 323
Db	765 TRYVKWAMATMMWSWRKMSYRMSGMBRWSAMRYCSMKCAKTYASASAMTKYAK 706
Oy	324 GTTACGATATATCCCGTTTCACATACCGGAAGTAGAAGCTTGGAATTAAGTTAA 383
Db	705 RSYRRRRWYKRGKWTYRYYRMSCBMTRAMSRRRKMAASMSKCMWYRGARSMTYS 646
Oy	384 CTCCTCGAGACAGTTAAGAGTGGGGGTACGACGGGATAGTGAACCGGAATGTCG 443
Db	645 KYSGSAKCKTKRMYTSSYSTGMTGMYSSYKSSWTSKMSYMGAMTCTMTTSMKGSRR 586
Oy	444 GTTTGGGTTTCAACATCACCATTAAGAAATAGACGATCCATCTGATGAGTGGTTGA 503
Db	585 SKMRMSMSMSMYRWRMKRKRKMYRMKCKTMRBMCYTRMGTYTYSRSMMYTR 526
Oy	504 TGGCAATATTCATTTGGCAACTTAAACGGGAAAGTTTATCAGCCGATTCATCTGCTG 562
Db	525 YKAYTTSRRRTMYKRYKCYYYYGMYMKSSYMMRYGYCAKCKCCYAMCMKAAYSMM 466
Oy	563 AGATATATGCAAAATCAAAATTCAGGACATATTCGACCGACCTCAACCCCTGCTGCCA 622
Db	465 MYTRTKTSKMRMSSTKTYMSMYKKCSMKYGAAGCTGCKRMYYCSGYMAYTYMSYK 406
Oy	623 TGACGACAAATCTGATTTGATATGCGGATTTTATGACATATGTCACCTCGGAGAC 682
Db	405 YSRCTKWBMYMYGMMYMYTYSAYSSMMTYTYYTAAKWKYVYRBRGMSYGSYKKY 346
Oy	683 ATTACTAC 691
Db	345 CTWMCYMKC 337

DEFINITION	Gorilla gorilla gorilla clone CH255-154N17, WORKING DRAFT SEQUENCE, 2 ordered pieces.
ACCESSION	AC144882
VERSION	AC144882.2 GI:31860084
KEYWORDS	HMG; HTGS PHASES; HTGS DRAFT.
SOURCE	Gorilla gorilla gorilla (lowland gorilla)
ORGANISM	Gorilla gorilla Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominae; Gorilla.
REFERENCE	1 (bases 1 to 204839) Antoneillis,A., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bonfield,G.G., Brinkley,C., Brooks,S., Cariaga,K., Chung,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X., Gupta,J., Haghghighi,P., Han,D., Hansen,N., Ho,S.-L., Hu,P., Hartle,B., Idol,J.R., Karlins,E., Khong,P., Laric,P., Lee-Jin,S.-O., Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Reddix-Dugue,N., Schandler,K., Scheuler,M.G., Shah,K., Slason,C., Stancirpop,S., Thomas,J.W., Thomas,P.D., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D. NSC Comparative Sequencing Initiative
TITLE	Unpublished
JOURNAL	2 (bases 1 to 204839) Green,E.D. Direct Submission Submitted (23-MAY-2003) NIH Intramural Sequencing Center, 8717 Grovermont Circle, Gathersburg, MD 20877, USA 3 (bases 1 to 204839) Green,E.D. Direct Submission Submitted (18-JUN-2003) NIH Intramural Sequencing Center, 8717 Grovermont Circle, Gathersburg, MD 20877, USA On Jun 18, 2003 this sequence version replaced gi:31044297.
COMMENT	-----Genome Center Center: NIH Intramural Sequencing Center Center code: NISC Web site: http://www.nisc.nih.gov Contact: nisc.zoo@nhgri.nih.gov -----Project Information Center project name: enf Center clone name: 154N17

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

-----Summary Statistics

```
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 204590 bases at least Q40
Consensus quality: 204680 bases at least Q30
Consensus quality: 204709 bases at least Q20
Insert size: 192000; agarose-fp
Insert size: 204739; sum-of-contigs
Quality coverage: 11.99x in Q20 bases; agarose-fp
Quality coverage: 11.24x in Q20 bases; sum-of-contigs
-----
```

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have * provided by the submittor.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP, Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep/Clone-derived-zebrafish_PUC_subclones_occasionally_display_inconsistency_over_the_length_of_mononucleotide_A/T_runs_and_conserved_TA_repeats.Where_this_is_found_the_longest_good_quality_representation_will_be_submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhixiong Rao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, Mashu). For further information see http://www.sanger.ac.uk/Projects/D_reio/fishmask.shtml DKEX-228C11 is from a zebrafish BAC library VECTOR: pindigobac-5.

FEATURES
source
1.163151
Location/Qualifiers

/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEX-228C11"
/clone_1ib="DantolKey"

ORIGIN

Query Match 5.2%; Score 38.6; DB 5; Length 163151;
Best Local Similarity 46.7%; Pred. No. 3.6;
Matches 122; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

5 GATTATGAAAGTTTGGATGATTTGGTACCGTATATCATGCTTTATTCAGTGGG 64
139459 GAAAGATGATATTTTGGGCTGATGCTTGGCATTGTCTTATCCGATATCCATTGGG 139518

65 GCGCAATGACAGATTTGCGGTTTGGCGGTTATGTTGGGCAATTAATGCTTCGCGCA 124
139519 AAGGTGCTATCCAGATGAGACCTGTCGACGAACGCTGCTCACTTAATGTTAGTTTTCG 139578

125 AACAGGTATATGATTCAGAGAAACACCGCGAAGCATCAGCGTAAAGCGCTACAA 184
139579 AATATCTGCTTATTCACAGCGGTGACAAATACCTGTAACCATTAATTAATATATAT 139638

185 CTCTTGTGTTATTCAGCAATATGTTACAGTATGACAGTAAGCAAGCAAACTGAGTA 244
139639 AATATGTTATCAATTTTATTTTATGACAAAACAACTCTGTAACCATGCGATGTC 139698

245 CGTCTCCGCTCATATGCTTA 265
139699 ATTCCCTCAGCTGMAATGCTTA 139719

RESULT 10
AC123226/c 244259 bp DNA linear HTG 08-OCT-2002
LOCUS
DEFINITION
*** 8 unordered pieces.

AC123226
AC123226.3 GI:23195011
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.

VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus clone CH230-217M18, *** SEQUENCING IN PROGRESS

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

AC123226
Rattus norvegicus (Norway rat)

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,

Fernandez, S., Fliley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Georgiev, E., Geor, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Guarine, E., Haaland, W., Hamill, C., Hamilton, N., Hernandez, J.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,

Hollins, B., Howells, S., Hui, J., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A.,

Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kovis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorenshewa, L., Louised, H., Lozada, R.J., Lu, X., Ma, J.,

Malleshwari, M., Mahindarc, M., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,

Manning, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, B., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L.,

Nankervil, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwackeleme, O., Okwou, G., Olampunagoon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,

Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L.,

Puaro, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,

Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,

Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,

Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smjs, D.,

Sneed, A., Sodergren, B., Song, X.-Z., Sorelle, R., Soza, J.,

Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, P., Uman, K.,

Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,

Wang, Q., Wang, S., Warren, J., Warren, J., Wei, X., White, P.,

Williams, G., Willis, R., Wleczyk, R., Wooden, H., Worley, K.,

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von

Weinert, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,

Wetters, G., and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 244259)

Worley, K.C.

TITLE

REFERENCE

AUTHORS

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

TITLE
JOURNAL
REFERENCE
AUTHORS
Worley, K.C.
DIRECT SUBMISSION
Submitted (09-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 257204)
Rat Genome Sequencing Consortium.
Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Oct 11, 2002 this sequence version replaced gi:21739166.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRIS
Center clone name: CH230-230B23
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 231067 bases at least Q40
Consensus quality: 235027 bases at least Q30
Consensus quality: 237586 bases at least Q20
Estimated insert size: 229826; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
2601: contig of 2601 bp in length
*
2602 2701: gap of unknown length
*
2702 232156: contig of 229455 bp in length
*
232157 232256: gap of unknown length
*
232257 241482: contig of 9226 bp in length
*
241483 241582: gap of unknown length
*
241583 242755: contig of 1173 bp in length
*
242756 242855: gap of unknown length
*
242856 243963: contig of 1108 bp in length
*
243964 244063: gap of unknown length
*
244064 245276: contig of 1213 bp in length
*
245277 245376: gap of unknown length
*
245377 246513: contig of 1137 bp in length
*
246514 246613: gap of unknown length
*
246614 248232: contig of 1619 bp in length
*
248233 248332: gap of unknown length
*
248333 249982: contig of 1650 bp in length
*
249983 250082: gap of unknown length
*
250083 252270: contig of 2188 bp in length

252271 252370: gap of unknown length
*
252371 253471: contig of 1101 bp in length
*
253472 253571: gap of unknown length
*
253572 255719: contig of 2148 bp in length
*
255720 255819: gap of unknown length
*
255820 257204: contig of 1385 bp in length.
FEATURES
source
1.257204
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-230B23"
1..1057
/note="wgs_end_extension
clone_end:5p6"
misc_feature
2702..2701
/estimated_length=unknown
2702..3930
/note="wgs_end_extension
clone_end:5p6"
misc_feature
3251..3632
/note="clone_boundary
clone_end:5p6
site:ECORI
220094..220968
/note="clone_boundary
clone_end:T7
site:ECORI
end_sequence:RMB0G12TV"
misc_feature
230583..232156
/note="wgs_end_extension
clone_end:T7"
gap
232157..232256
/estimated_length=unknown
241483..241582
/estimated_length=unknown
242756..242855
/estimated_length=unknown
243964..244063
/estimated_length=unknown
245277..245376
/estimated_length=unknown
246514..246613
/estimated_length=unknown
248233..248332
/estimated_length=unknown
249983..250082
/estimated_length=unknown
252271..252370
/estimated_length=unknown
253472..253571
/estimated_length=unknown
255720..255819
/estimated_length=unknown
gap
/estimated_length=unknown
ORIGIN
Query Match 5.2%; Score 38.6; DB 14; Length 257204;
Best Local Similarity 54.6%; Pred. No.3.7;
Matches 77; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 504 TGCGATATTCATTGCGAAGCTTAACGGGACGTTATACGCCGATTCATGCTGA 563
DB 28833 TGGCATTTCTGGAATTCAGAGCGCTTATATGTAAGAGTAGCCATCTGAACCTACTACCTTA 28892
QY 564 GATATATGCAATATGAATTTCAGAGCATATCCGACCGACCTCAACCTCGTGTTCAT 623
DB 28893 CATTTTCCAGAGCCATTCATACATGAGCCATCACCCCAACCCATATATGTATAT 28952
QY 624 GACGACAATCTCGTATTTGA 644
DB 28953 AGTTCTAAGACCTACATTGA 28973

RESULT 12
BX571959/c
LOCUS
DEFINITION BX571959 116374 bp DNA linear VRT 05-NOV-2004
Zebrafish DNA sequence from clone DKEX-4616 in linkage group 12,
complete sequence.
ACCESSION BX571959
VERSION BX571959.8 GI:55467255
KEYWORDS HTG.
SOURCE Dantio rerio (zebrafish)
ORGANISM Dantio rerio
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 116374)
REFERENCE
AUTHORS Barker, D.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 5, 2004 this sequence version replaced gi:54304257.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMBPP; Information
on the WORMBPP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpe/Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhirong Bao and Sean Rddy, submitted), and those
beginning 'dir' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEX-4616
is from a Zebrafish BAC library
VECTOR: pindigobac-5.
FEATURES
source
1..116374
/organism="Dantio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEX-4616"
/clone_1lb="DantioKey"
ORIGIN
Query Match 5.1%; Score 38.2; DB 5; Length 116374;
Best Local Similarity 59.8%; Pred. No. 4.8;
Matches 64; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 508 AATTCATTTGCAAGACTTAAACGGGACGTTTATCAGCCGAGTTCTACTGCTAGATA 567
DB 47200 AATTCCTAATTTAGTTCTTACGACGAAATTTACTTACCAATATTAAGTTTGTATATA 47141

QY 568 TATGCAATATCAATTTCAGGACTATTCGACCGACTCAACCTCG 614
DB 47140 TTTTCTGTAGAAATTCACAGATATTTTCTATAGAACTCAACCTTG 47094
RESULT 13
CR933779/c
LOCUS
DEFINITION CR933779 242081 bp DNA linear HTG 10-JUN-2005
Dantio rerio chromosome 12 clone DKEX-1305, WORKING DRAFT SEQUENCE,
12 unordered pieces.
ACCESSION CR933779
VERSION CR933779.4 GI:67509338
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS _FULLTOP.
SOURCE Dantio rerio (zebrafish)
ORGANISM Dantio rerio
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 242081)
REFERENCE
AUTHORS McLaren, S.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests:
http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeighc
On Jun 10, 2005 this sequence version replaced gi:58330714.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: zK1305
----- Summary Statistics
Assembly program: XGAP; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 236648 bases at least Q40
Consensus quality: 237713 bases at least Q30
Consensus quality: 238743 bases at least Q20
Insert size: 240981; sum-of-contigs
Insert size: 168181; 37.4% error; agarose-fp
Quality coverage: 6.79x in Q20 bases; sum-of-contigs Quality
coverage: 1.84x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
3578: contig of 3578 bp in length
1
3579 3578: gap of 100 bp
* 3579 3578: contig of 1008 bp in length
* 13687 13786: gap of 100 bp
* 13687 45507: contig of 31721 bp in length
* 45508 45607: gap of 100 bp
* 45608 60434: contig of 14827 bp in length
* 60435 60534: gap of 100 bp
* 60535 72346: contig of 11812 bp in length
* 72347 72447: gap of 100 bp
* 72447 79048: contig of 6602 bp in length
* 79049 79149: gap of 100 bp
* 79149 90278: contig of 11130 bp in length
* 90279 90378: gap of 100 bp
* 90379 131546: contig of 41168 bp in length
* 131547 131646: gap of 100 bp
* 131647 141511: contig of 965 bp in length
* 141512 141711: gap of 100 bp
* 141712 173884: contig of 32173 bp in length
* 173885 173984: gap of 100 bp

* 173985 237730: contig of 63746 bp in length
* 237731 237830: gap of 100 bp
* 237831 242081: contig of 4251 bp in length.

FEATURES
source
1. 242081
/organism="Dantio xerico"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/chromosome="12"
/clone_1ib="DantioKey"
1. .3578
/note="assembly fragment:01599
fragment chain:1"
misc_feature
3679. .13686
/note="assembly fragment:01661
fragment chain:1"
misc_feature
13787. .45507
/note="assembly fragment:02271
fragment chain:1"
misc_feature
45608. .60434
/note="assembly fragment:02110
fragment chain:1"
misc_feature
60535. .72346
/note="assembly fragment:01984
fragment chain:1"
misc_feature
72447. .79048
/note="assembly fragment:01793
fragment chain:1"
misc_feature
79149. .90278
/note="assembly fragment:01867
fragment chain:1"
misc_feature
90379. .131546
/note="assembly fragment:00001
fragment chain:1"
misc_feature
131647. .141611
/note="assembly fragment:01720
fragment chain:2"
misc_feature
141712. .173884
/note="assembly fragment:00406
fragment chain:2"
misc_feature
173985. .237730
/note="assembly fragment:00838.0"
misc_feature
237831. .242081
/note="assembly fragment:01615"
ORIGIN
Query Match 5.1%; Score 38.2; DB 14; Length 242081;
Best Local Similarity 59.8%; Pred. No. 4.9; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 508 ATATTCATTGTCAGAACTTAACGGGAACGTTATCACCGGATTTCTACTGCTGAGATA 567
|||||
Db 220878 ATATTCGAATTTAGTTCTTGAAGAGAAATTTACTAACAGATATTAAGTTTGTATTA 220819
|||||

QY 568 TATGCAATTCAGAACTTATTCGACCGACCTCAACCTCG 614
|||||
Db 220818 TTTTCGTGAATTCACAGATATTTTCATAGAACTCAACCTTGG 220772
|||||

RESULT 14
AF225410 576 bp mRNA 1linear PLN 21-JUL-2005
DEFINITION Gastrodia elata antifungal protein precursor, mRNA, complete cds.
ACCESSION AF225410
VERSION AF225410.2 GI:71040863
KEYWORDS
SOURCE
ORGANISM
Gastrodia elata
Gastrodia elata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
Epilendroidae; lower Epilendroidae; Gastrodiae; Gastrodia.

REFERENCE
1 (bases 1 to 576)

AUTHORS
TITLE
Wang, Y., Li, W. and Guo, S.
Purification, in vitro Activity and cDNA Cloning of Gastrodia
Antifungal Protein (GAPP) from Gastrodia (Gastrodia elata)

JOURNAL
TITLE
Unpublished
2 (bases 1 to 576)
AUTHORS
Wang, Y., Li, W., Fu, R., Guo, S. and Sun, Y.
DIRECT SUBMISSION
Submitted (14-JAN-2000) Genetics, Institute of Genetics, Datun
Road, Beijing 100101, China
3 (bases 1 to 576)
AUTHORS
Wang, Y., Li, W., Fu, R., Guo, S. and Sun, Y.
DIRECT SUBMISSION
Submitted (20-JUL-2005) Genetics, Institute of Genetics, Datun
Road, Beijing 100101, China
SEQUENCE UPDATE BY SUBMITTER
On Jul 20, 2005 this sequence version replaced gi:12330253.

REMARK
COMMENT
On Jul 20, 2005 this sequence version replaced gi:12330253.

FEATURES
source
1. .576
/organism="Gastrodia elata"
/mol_type="mRNA"
/db_xref="taxon:91201"
/tissue_type="secondary corn"
1. .540
/codon_start=1
/product="antifungal protein precursor"
/protein_id="AG52864.2"
/db_xref="GI:71040864"
/translation="MAASASTAVILFFAVTTMMSSALPAPASDRINGHOLDTGSGL
AAGGFLTIQNDGNTLVLDNDRVAVASGNGASGCVLMQNDGNLVYSGRAIVAS
NTNRQNGVYLLIDRDHNVIVYDNNNMIAWHTHVGNALIPAHSNTAASGAAG
NKVNLVISMYSRSRIAG"
1. .84
1. .537
/product="antifungal protein"

ORIGIN
Query Match 5.1%; Score 37.8; DB 15; Length 576;
Best Local Similarity 58.4%; Pred. No. 5.2; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 379 GTTAACCTCCCTGACCAAGTAAAGATGTCGGGTACGAGGATGTAACCGCAAT 438
|||||
Db 325 GGTAACTACTATCTATCTTACGAGAGATGTAACGTCGTCTATATCAATATTTAAT 384
|||||

QY 439 GGTGCGTTTGGGTATCAACATCAACATAGAAATGACGATCCATCTCTAT 491
|||||
Db 385 AATGCGATTGGGCAACCAACCAACGTTGAAATGCTGAATATCATCTCCAT 437
|||||

RESULT 15
AY032588 1554 bp DNA 1linear PLN 11-JUN-2001
DEFINITION Gastrodia elata antifungal protein gene, complete cds.
ACCESSION AY032588
VERSION AY032588.1 GI:14346011
KEYWORDS
SOURCE
ORGANISM
Gastrodia elata
Gastrodia elata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
Epilendroidae; lower Epilendroidae; Gastrodiae; Gastrodia.

REFERENCE
1 (bases 1 to 1554)
AUTHORS
Sun, Y.-R., Wang, Y.-Q., Li, W.-B., Zhang, L.-M., Niu, H.-Y. and
Sun, Y.-R.
Gastrodia elata antifungal protein gene
Unpublished
2 (bases 1 to 1554)
AUTHORS
Sun, Y.-R., Wang, Y.-Q., Li, W.-B., Zhang, L.-M., Niu, H.-Y. and
Sun, Y.-R.
DIRECT SUBMISSION
Submitted (15-APR-2001) Group 603, Institute of Genetics, Da Tun
Road, Beijing 100101, People's Republic of China
LOCATION/Qualifiers

FEATURES
TITLE
JOURNAL
JOURNAL
Submitted (15-APR-2001) Group 603, Institute of Genetics, Da Tun
Road, Beijing 100101, People's Republic of China
LOCATION/Qualifiers

source 1..1554
/organism="Gastrodia elata"
/mol_type="genomic DNA"
/db_xref="taxon:91201"
misc_feature 1..581
/note="contains promoter and 5' UTR"
<582..>1091
/product="antifungal protein"
582..1091
CDS
/codon_start=1
/product="antifungal protein"
/protein_id="AK59994.1"
/db_xref="GI:14346012"

translation="MASPASAVIFLPVPAALMSLAMPALAAASOLNAGOTLGTGSL
AOGPDQVIONDNVLVDNRRVWVWASGTNGKAGSCVLRMORDGRLVYSGSRVTWAS
NTNRRDNYILLQDRRVVVIYDSSNNAIWAIGTVNGNAITVIFHSNGTAAAGAAQ
NKVNEYLRP"

ORIGIN

Query Match 5.1%; Score 37.8; DB 15; Length 1554;
Best Local Similarity 58.4%; Pred. No. 5.4;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 379 GTTAACCTCCCTGCACCACTTAAGAGTGTGGGTTACGACGGGATAGTGAAACCGCAAT 438
DB 906 GATACTACTATCTGCTCTCCCTCAGAGAGACGTAACGTTGTCATATACGATCTTCAAT 965
QY 439 GGTGGTTTGGGTATCAACATCAACATAGAAATGACGGATCCACTCTTAT 491
DB 966 AATGCAATTGGGCAACCGGCACCAACGTTGAAATGCTGCAATCACTGTAT 1018

Search completed: April 7, 2006, 22:44:12
Job time : 4222 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 7, 2006, 21:06:31 ; Search time 587 Seconds
(without alignments)
8447.242 Million cell updates/sec

Title: US-10-784-592-18

Sequence: 1 gtgcgaatcgtgaaagttt.....aaacgacgtatgctcttcg 744

Scoring table: IDENTITY_NUC

Searched: 4996997 beqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

```

1:  _  geneseqn1980s:*
2:  geneseqn1990s:*
3:  geneseqn2000s:*
4:  geneseqn2001as:*
5:  geneseqn2001as:*
6:  geneseqn2002as:*
7:  geneseqn2002as:*
8:  geneseqn2003as:*
9:  geneseqn2003as:*
10: geneseqn2003cs:*
11: geneseqn2003as:*
12: geneseqn2004as:*
13: geneseqn2004as:*
14: geneseqn2005as:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	744	100.0	744	14	ABE45544	AbE45544 Allicyclob	
2	744	100.0	744	14	ABE48730	AbE48730 Allicyclob	
3	621	83.5	621	14	ABE45582	AbE45582 Allicyclob	
C	39.4	5.3	2000	8	ADA71938	Ada71938 Rice gene	
	36.6	4.9	2000	8	ADA71938	Ada71938 Rice gene	
	36.2	4.9	2306	8	ACC47595	Acc47595 Gastridia	
	36.2	4.9	110000	9	ADA13411_1	Continuation (2 of	
7	36.2	4.9	110000	9	ADOS59445_1	Continuation (2 of	
8	36.2	4.9	110000	12	ADZ13757_1	Continuation (2 of	
9	36.2	4.9	110000	14	ADZ13757_1	Continuation (2 of	
10	35.6	4.8	7636	5	ABA17217	AbA17217 Human ner	
C	11	35.2	4.7	580	10	ADK56798	AdK56798 Plant DNA
	11	35	4.7	2349	13	ADK561627	AdK561627 Bacterial
	13	35	4.7	42488	4	AAK6772	AAK6772 Human imm
	13	34	4.6	2000	11	ACL37108	ACL37108 Rice stre
14	34	4.6	2000	11	ACL37108	ACL37108 Rice stre	
15	33.8	4.5	729	13	ADX12170	AdX12170 Plant full	
16	33.8	4.5	860	13	ADX12259	AdX12259 Plant full	
17	33.4	4.5	706	10	ADK56319	AdK56319 Plant DNA	
18	33.4	4.5	1572	3	AAZ53068	AAZ53068 Neisseria	
19	33.4	4.5	1848	10	ABZ38315	AbZ38315 N. gonorr	

20	33.4	4.5	1851	3	AAZ53069	AAZ53069 <i>Neisseria</i>
C 21	33.4	4.5	12077	3	AAAB1734	AAAB1734 <i>N. mening</i>
C 22	33.4	4.5	11000	3	AAAB1489_6	Continuation (7 of
C 23	33.4	4.5	34980	3	AAAF21612	AAAF21612 <i>Neisseria</i>
C 24	33.2	4.5	7480	13	ADP25286	ADP25286 <i>Pro polyA</i>
C 25	33.2	4.5	7480	14	ADY17856	ADY17856 <i>DNA encod</i>
C 26	32.6	4.4	2700	5	AAD14385	AAD14385 <i>DNA encod</i>
C 27	32.4	4.4	813	8	ACA28717	ACA28717 <i>Mouse bon</i>
C 28	32.4	4.4	18595	2	AAAS33411	AAAS33411 <i>Prokaryot</i>
C 29	32.4	4.4	49999	2	AAZ23869	AAZ23869 <i>DNA encod</i>
C 30	32.4	4.4	301477	13	ABD33362	ABD33362 <i>Human LOB</i>
C 31	32.2	4.3	123526	10	ADJ79962	ADJ79962 <i>Human</i>
C 32	32	4.3	600	2	AAV86185	AAV86185 <i>EST clone</i>
C 33	32	4.3	26147	4	ABL11324	ABL11324 <i>Drosophila</i>
C 34	31.8	4.3	503	6	ABO50691	ABO50691 <i>Oligonuc1</i>
C 35	31.8	4.3	503	6	ABO50690	ABO50690 <i>Oligonuc1</i>
C 36	31.8	4.3	534	6	ABO55323	ABO55323 <i>Human ova</i>
C 37	31.8	4.3	1317	4	AAAF61101	AAAF61101 <i>P. putida</i>
C 38	31.8	4.3	83432	14	ABEA61124_3	Continuation (4 of 5)
C 39	31.8	4.3	110000	14	ADZ45062_07	Continuation (8 of 6)
C 40	31.6	4.2	633	4	AA822538	AA822538 <i>Human CDN</i>
C 41	31.6	4.2	634	4	AA822774	AA822774 <i>Human CDN</i>
C 42	31.6	4.2	704	10	ADK58633	ADK58633 <i>Plant DNA</i>
C 43	31.6	4.2	1007	10	ADK56823	ADK56823 <i>Plant DNA</i>
C 44	31.6	4.2	2000	11	ACL35887	ACL35887 <i>Rice stre</i>
C 45	31.6	4.2	110000	10	ACPF67367_48	Continuation (49 of

ALIGNMENTS

```

RESULT 1
ID AEB45544 standard; DNA; 744 BP.
XX AEB45544;
AC AEB45544;
XX
D7 22-SEP-2005 (first entry)
XX
DE Alicyclobacillus sp. functional polypeptide DNA, SEQ ID NO: 18.
XX
KW Feedstuff; food; detergent; surfactant; pulp; functional polypeptide;
   gene; ds.
XX
OS Alicyclobacillus sp.; DSM 15716.
XX
XX
Key Location/Qualifiers
CDS 1..744
   FT 1..744
   FT /*tag= b
   FT /product= "Alicyclobacillus sp. functional polypeptide"
   FT /transl_except= (pos:1..3, aa:Mec)
   FT /partial
   FT /note= "No start and stop codons"
   FT 1..123
   FT /*tag= a
   FT 124..744
   FT /*tag= c
   FT /product= "Alicyclobacillus sp. mature functional
   FT polypeptide"
   PN US2005147983-A1.
   PD 07-JUL-2005.
   PD 23-FEB-2004; 2004US-00784592.
   PD 06-JAN-2004; 2004DK-00000010.
   PD 04-FEB-2004; 2004DK-00000165.
   PA (NOVO ) NOVOZYMES AS.
   PI Wilting R, Lassen SF, Ostergaard PR;
   XX

```

DR WP1: 2005-511773/52.
 DR P-PSDB; AEB45569.
 PT New functional polypeptides having function and amino acid sequence
 PT similar to known specific bacterial enzymes useful in industrial,
 PT research and household applications e.g. detergents and food.
 PS Example 3; SEQ ID NO 18; 83bp; English.
 XX
 XX The present invention relates to functional polypeptides encoded by
 CC polynucleotides comprised in the genome of *Allicyclobacillus* sp. DSM
 CC 15716. The invention is useful in industrial, research and household
 CC processes such as in composition to prepare feed and food, in detergent
 CC formulations and for treating lignocellulosic fabric and pulp. The
 CC present sequence is *Allicyclobacillus* sp. DSM 15716 functional polypeptide
 CC DNA.
 XX
 XX Sequence 744 BP; 202 A; 159 C; 179 G; 204 T; 0 U; 0 Other;

Query Match 100.0%; Score 744; DB 14; Length 744;
 Best Local Similarity 100.0%; Pred. No. 5,7e-237;
 Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAATTATGAAAGTTTGGATGATTTGTACCGTATATCATGCTTTTATTCAG 60
 DB 1 GTGCAATTATGAAAGTTTGGATGATTTGTACCGTATATCATGCTTTTATTCAG 60
 QY 61 TGGGGGCGAATGAAAGAAATCTGCGTTTGGCGGTTTCATTGTGGGCAATTATGCTTC 120
 DB 61 TGGGGGCGAATGAAAGAAATCTGCGTTTGGCGGTTTCATTGTGGGCAATTATGCTTC 120
 QY 121 GCGAACACGGTGTATATGATTCGAGAAACACACCGCGGAAAGCATCAACGGTAACGCT 180
 DB 121 GCGAACACGGTGTATATGATTCGAGAAACACACCGCGGAAAGCATCAACGGTAACGCT 180
 QY 181 ACAACTTCTTGGTTAATTGACGAAATGTTCAAGGTAGCAAGCAAGCAAAATCTCG 240
 DB 181 ACAACTTCTTGGTTAATTGACGAAATGTTCAAGGTAGCAAGCAAGCAAAATCTCG 240
 QY 241 AGTAGTCTCCCGCTCATAGTCTAGCAATCTGCAACATGCAACATGCAACATGCAAC 300
 DB 241 AGTAGTCTCCCGCTCATAGTCTAGCAATCTGCAACATGCAACATGCAACATGCAAC 300
 QY 301 ACGACTTATCTTCTCACTGCGAAGTTACGATATATCCGTTTCAACATACGGGAAGGTA 360
 DB 301 ACGACTTATCTTCTCACTGCGAAGTTACGATATATCCGTTTCAACATACGGGAAGGTA 360
 QY 361 GGAAGCTTGAATTAAGTTAACTCCGTCAGCAAGTTAAGAGTGGGGTACGACGGG 420
 DB 361 GGAAGCTTGAATTAAGTTAACTCCGTCAGCAAGTTAAGAGTGGGGTACGACGGG 420
 QY 421 ATAGGTGAACCGGCAATGCGGTTGGGTTATCAACATCAACATGAAGAAATACCGGA 480
 DB 421 ATAGGTGAACCGGCAATGCGGTTGGGTTATCAACATCAACATGAAGAAATACCGGA 480
 QY 481 TCCACTCTATGAGAGTGTGATGAGCATATTCATTTGACAGACTTAAACGGGAAGCT 540
 DB 481 TCCACTCTATGAGAGTGTGATGAGCATATTCATTTGACAGACTTAAACGGGAAGCT 540
 QY 541 TATCAGCCGGATTTACTGCTGAGATATATGCAAAATACAAATTCAGGAGCTATTCGAC 600
 DB 541 TATCAGCCGGATTTACTGCTGAGATATATGCAAAATACAAATTCAGGAGCTATTCGAC 600
 QY 601 GACCTCAACCCCTGGTGTGTCATGACGAATAATCTCGATATTTGAATGCGGATTTTATG 660
 DB 601 GACCTCAACCCCTGGTGTGTCATGACGAATAATCTCGATATTTGAATGCGGATTTTATG 660
 QY 661 ACATATGTCAGCTGCGGAGCATTAATCACTTGTGCTTCCATGAGGTTTCTTCGGGTCA 720
 DB 661 ACATATGTCAGCTGCGGAGCATTAATCACTTGTGCTTCCATGAGGTTTCTTCGGGTCA 720
 QY 721 GATGAAACGACGTATGCTTCTCG 744
 DB 721 GATGAAACGACGTATGCTTCTCG 744

DB 721 GATGAAACGACGTATGCTTCTCG 744
 RESULT 2
 ID AEB48730 standard; DNA; 744 BP.
 AC AEB48730/
 XX
 XX 22-SEP-2005 (first entry)
 DE *Allicyclobacillus* sp. DSM 15716 functional polypeptide coding sequence.
 KW Protein production; protein secretion; surfactant; feedstuff; food; gene;
 ds.
 XX
 XX *Allicyclobacillus* sp.; 'DSM 15716'.
 OS
 XX
 XX Key Location/Qualifiers
 FH 1..744
 FT CDS /tag= b
 FT /transl_except= (pos:1..3,aa:Met)
 FT /partial
 FT sig_peptide /note= "No stop codon"
 FT 1..123
 FT /tag= a
 FT 124..744
 FT mat_peptide /tag= c
 FT /product= "Mature polypeptide"
 PN
 XX
 XX WO200506339-A2.
 PD
 XX
 XX 21-JUL-2005.
 PF
 XX
 XX 06-JAN-2005; 2005WO-DK000004.
 PR
 XX
 XX 06-JAN-2004; 2004DK-00000010.
 PR
 XX
 XX 04-FEB-2004; 2004DK-00000165.
 PR
 XX
 XX 23-FEB-2004; 2004US-00784592.
 PR
 XX
 XX 25-FEB-2004; 2004DK-00000298.
 PA
 XX
 XX (NOVO) NOVOZYMES AS.
 PI
 XX
 XX Wiltling R, Lassen SF, Ostergaard PR;
 DR
 XX
 XX WP1: 2005-506869/51.
 DR
 XX
 XX P-PSDB; AEB48755.
 PT
 XX
 XX New mature functional polypeptide of *Allicyclobacillus* sp., used for
 PT commercial research purposes, specifically for cleaning a cellulosic
 PT fabric, preparing food or feed additive, and for treating lignocellulosic
 PT materials and pulp.
 PS
 XX
 XX Claim 3; SEQ ID NO 18; 151bp; English.
 XX
 XX The present invention is based on the finding of a strain of
 CC *Allicyclobacillus*, namely *Allicyclobacillus* sp. DSM 15716, which grows at
 CC low pH (approximately 4-5) and at high temperature (50-60 degrees C). It
 CC is an object of the invention to identify and provide polypeptides
 CC secreted from *Allicyclobacillus* sp. DSM 15716 because such polypeptides
 CC may be used for industrial purposes and may also be produced in
 CC industrially relevant processes and amounts. Thus, the invention provides
 CC isolated mature functional polypeptides which are at least 90% identical
 CC to, and exhibit and same function as a corresponding secreted polypeptide
 CC AEB48758-AEB48762 from *Allicyclobacillus* sp. DSM 15716, and also provides
 CC polynucleotides AEB48730-AEB48737 encoding such polypeptides. The
 CC polynucleotides are preferably enzymes having acid endoglucanase, acid
 CC cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl
 CC protease, serine protease, HtrA-like serine protease, disulfide
 CC isomerase, gamma-D-glutamyl-L-amino acid endopeptidase, endo-beta-N-
 CC acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase,
 CC phylase, phospholipase C, polysaccharide deacetylase, xylanase
 CC deacetylase or sulfite oxidase activity. A composition comprising such an

enzyme can be used in a detergent composition, or a food or feed composition. Also claimed are: nucleic acid constructs, recombinant expression vectors and host cells, methods of producing the polypeptides; a storage medium suitable for use in an electronic device comprising information of the amino acid sequence of a polypeptide of the invention or the nucleotide sequence encoding the polypeptide, and a process comprising employing the polypeptide, or polynucleotide encoding it, in an industrial or household technical process. The present sequence is the coding sequence of a functional polypeptide of the invention.

Sequence 744 BP; 202 A; 159 C; 179 G; 204 T; 0 U; 0 Other;

Query Match 100.0%; Score 744; DB 14; Length 744;

Best Local Similarity 100.0%; Pred. No. 5.7e-237; Mismatches 0; Indels 0; Gaps 0;

Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGCGAATTATGAAAGTTTGGATGATTTGGTACCGTATATCAATGCTGTTATTCAG 60
DB 1 GTGCGAATTATGAAAGTTTGGATGATTTGGTACCGTATATCAATGCTGTTATTCAG 60
QY 61 TGGGGGCGGAATGACGAAATTCGCTTTGCCGTTCAATTGGGGCATTAAATGTCTTC 120
DB 61 TGGGGGCGGAATGACGAAATTCGCTTTGCCGTTCAATTGGGGCATTAAATGTCTTC 120
QY 121 GCGAACACGGTGTATGATTCGAGGAAACACCGCGGAGCATCAACGGTAAGCGCT 180
DB 121 GCGAACACGGTGTATGATTCGAGGAAACACCGCGGAGCATCAACGGTAAGCGCT 180
QY 181 ACAACTCTCTTGGTATTAATTCAGCATATGTTCAACAGTACGAAAGCAAGCAAACTCG 240
DB 181 ACAACTCTCTTGGTATTAATTCAGCATATGTTCAACAGTACGAAAGCAAGCAAACTCG 240
QY 241 AGTACGTCCTCCGCTCATATGCTACGAACTCATTCGCAATGCGCAATCATAGCTGCT 300
DB 241 AGTACGTCCTCCGCTCATATGCTACGAACTCATTCGCAATGCGCAATCATAGCTGCT 300
QY 301 ACGACTCTCATCTTCAGTCGAAAGTTAGATATATCCGTTTCAACATACCGGAAGGTA 360
DB 301 ACGACTCTCATCTTCAGTCGAAAGTTAGATATATCCGTTTCAACATACCGGAAGGTA 360
QY 361 GGAGACTTGGAAATTAAGTAACTCCCTGACGACAAATTAGAGTGTGGGTACGACGG 420
DB 361 GGAGACTTGGAAATTAAGTAACTCCCTGACGACAAATTAGAGTGTGGGTACGACGG 420
QY 421 ATAGGTGAACCGCAAAATGTCGCTTTGGGTTATCAACATCAACATTAAGAAATGACGA 480
DB 421 ATAGGTGAACCGCAAAATGTCGCTTTGGGTTATCAACATCAACATTAAGAAATGACGA 480
QY 481 TCCACTCCATGAGAGTGTGATGAGCAATTCATTGTCAGAACTTAAACGGGAACGTT 540
DB 481 TCCACTCCATGAGAGTGTGATGAGCAATTCATTGTCAGAACTTAAACGGGAACGTT 540
QY 541 TATCAGCCGGAATTTACTGCTGAGATATATGCAAAATTAATTCAGGAGCTATTCGACC 600
DB 541 TATCAGCCGGAATTTACTGCTGAGATATATGCAAAATTAATTCAGGAGCTATTCGACC 600
QY 601 GACCTCAACCCCTGTGTGTCCATGACGAACAATCTCGTATGATATGCGGATTTTATG 660
DB 601 GACCTCAACCCCTGTGTGTCCATGACGAACAATCTCGTATGATATGCGGATTTTATG 660
QY 661 ACATATGTCAGTGTGGGAGATATCACTTGTGCGTTCCATAGGTTTCTTCCGGTCA 720
DB 661 ACATATGTCAGTGTGGGAGATATCACTTGTGCGTTCCATAGGTTTCTTCCGGTCA 720
QY 721 GATGAACGAGCATGCTCTTCG 744
DB 721 GATGAACGAGCATGCTCTTCG 744
```

RESULT 3
AEB45582
ID AEB45582 standard; DNA; 621 BP.

```
AC AEB45582;
XX
DT 22-SEP-2005 (first entry)
XX
DE Alicyclobacillus sp. mature functional polypeptide DNA (bases 124-744).
XX
KW Feedstuff; food; detergent; surfactant; pulp; functional polypeptide;
KM gene; de.
XX
OS Alicyclobacillus sp.; DSM 15716.
XX
FH Key Location/Qualifiers
FT CDS 1..621
FT /tag= a
FT /product= "Alicyclobacillus sp. mature functional
FT polypeptide"
FT /partial
FT /note= "No start and stop codons"
XX
PN US2005147983-A1.
XX
XX 07-JUL-2005.
XX
PF 23-FEB-2004; 2004US-00784592.
XX
PR 06-JAN-2004; 2004DK-00000010.
XX 04-FEB-2004; 2004DK-00000165.
XX
PA (NOVO ) NOVOZYMES AS.
XX
PI Wiltling R, Lassen SF, Ostergaard PR.
XX
DR WPI; 2005-511773/52.
XX
DR P-PSDB; AEB45583.
XX
PT New functional polypeptides having function and amino acid sequence
PT similar to known specific bacterial enzymes useful in industrial,
PT research and household applications e.g. detergents and food.
XX
XX
PS Claim 36; Page; 83pp; English.
XX
CC The present invention relates to functional polypeptides encoded by
CC polynucleotides comprised in the genome of Alicyclobacillus sp. DSM
CC 15716. The invention is useful in industrial, research and household
CC processes such as in composition to prepare feed and food, in detergent
CC formulations and for treating lignocellulosic fabric and pulp. The
CC present sequence is Alicyclobacillus sp. DSM 15716 mature functional
CC polypeptide DNA. Note: This sequence is not shown in the specification
CC but is constructed based on the nucleotide positions provided in claim 36
CC of the specification.
XX
SQ Sequence 621 BP; 177 A; 142 C; 143 G; 159 T; 0 U; 0 Other;
XX
Query Match 83.5%; Score 621; DB 14; Length 621;
Best Local Similarity 100.0%; Pred. No. 5.4e-136; Mismatches 0; Indels 0; Gaps 0;
Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 124 AACACGGTATATGATTCGAGAAACACACCGCGGAAGCATCAACGGTAAGCGCTACA 183
DB 1 AACACGGTATATGATTCGAGAAACACACCGCGGAAGCATCAACGGTAAGCGCTACA 60
QY 184 ACTTTTGGTAAATTCGACGAATATGTTCAACAGTAGCAAGCAAGCAAAATCTGAGT 243
DB 184 ACTTTTGGTAAATTCGACGAATATGTTCAACAGTAGCAAGCAAGCAAAATCTGAGT 120
QY 244 ACGTCTCCGCTCATATGATTCGACGAATATGTTCAACAGTAGCAAGCAAGCAAAATCTGAGT 303
DB 244 ACGTCTCCGCTCATATGATTCGACGAATATGTTCAACAGTAGCAAGCAAGCAAAATCTGAGT 180
QY 304 ACTTCATCTTCTCAGTCGAAAGTTAGATATATCCCGTTTCAACATACGGAAGGTAGGA 363
DB 304 ACTTCATCTTCTCAGTCGAAAGTTAGATATATCCCGTTTCAACATACGGAAGGTAGGA 240
```

QY	364	GACCTTGGAAATTAGAGTTAACTCCCTCGACAGCAATTAAAGTGTGGGGTATCGACCGGGATA	423
Db	241	GACTTGGAAATTAGAGTTAACTCCCTCGACAGCAATTAAAGTGTGGGGTATCGACCGGGATA	300
QY	424	GGTGAACCCGCAAAATGSGTCGTTTGGGTTATCAACATCAACCATTAAGAAATGACGATCC	483
Db	301	GGTGAACCCGCAAAATGSGTCGTTTGGGTTATCAACATCAACCATTAAGAAATGACGATCC	360
QY	484	ACTCCTATGAGAGTCGTTGATGGCATATTCATCTTTGCAAGACTTAAACGGAAAGTTTAT	543
Db	361	ACTCCTATGAGAGTCGTTGATGGCATATTCATCTTTGCAAGACTTAAACGGAAAGTTTAT	420
QY	544	CAGCCGGGATCTACGTCGTGAGATATATGCAAAATACAAATTCAGGAGCTATTCGACCCGAC	603
Db	421	CAGCCGGGATCTACGTCGTGAGATATATGCAAAATACAAATTCAGGAGCTATTCGACCCGAC	480
QY	604	CTCAACCCCTGCGTGTCTCATGACGACAAATCTGTAATTGATATGCGCGAGTTTATATGACA	663
Db	481	CTCAACCCCTGCGTGTCTCATGACGACAAATCTGTAATTGATATGCGCGAGTTTATATGACA	540
QY	664	TATGTCACGTCGGGAGCATTTACTCATCTGTGCTTCACATGGGTTTCTTCGGGTCAGAT	723
Db	541	TATGTCACGTCGGGAGCATTTACTCATCTGTGCTTCACATGGGTTTCTTCGGGTCAGAT	600
QY	724	GAAAGCAGTATGCTTTCCG	744
Db	601	GAAAGCAGTATGCTTTCCG	621

RESULT 4
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.

AC ADA71938;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

KM Plant; bacterial infection; fungal infection; viral infection; rice, KM gene; ds.

OS *Oryza sativa*.

PN WO2003000898-A1.

PD 03-JAN-2003.

PF 22-JUN-2001; 2001WO-IB001105.

PR 22-JUN-2001; 2001WO-1B001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Chang H, Chen W, Cooper B, Glazebrook J, Golf SA, Hou Y;

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match	5.3%	Score 39.4;	DB 8;	Length 2000;
Best Local Similarity	9.1%	Pred. No. 0.086;		
Matches 61; Conservative	304;	Mismatches 300;	Indels 4;	Gaps 2

QY	24	ATGATTTTGGACCGGTATATCATGCTGTTATTAATGAGGGGGGCAATGAACGAAATTC	83
Db	1002	MMKTKTKMTATSSWTKMYTAAYMKRAYAMSRKTKWCTGGKBAATYTCGTMMMAAGRWK	943
QY	84	GGCTTTTCCGGTTCATTGTGGGCATTAATGTCTTCGCGAAACACGGTGTATATGATTCG	143
Db	942	BNAMCMYCCMMKMMKMTCCMMKKYRTWSCWITMMGAMRYAYYAMRBRRTMYTKMSWRM	883
QY	144	AGGAACACACCGCGGAACGATCAACGGTAAACGCTCAACTCTTGTGTTAATTCGAC	203
Db	882	YWTMKMAMTWTCOMAKMTATGATGMATMMWRMYTYCYAMTCAKCKTKMAATKMWT	823
QY	204	GAATAGTTCACAGTAGGACAAAGCAAGCAAAAACCTCGATAGCTCCCGCTCAATAGTC	263
Db	822	WACABRATSWRBRBAVAGMRKRYKMKRAYMWRMWRMCPVAGNARMKSYR--WIKKKYA	766
QY	264	TACGAATCATTCGCAACATCGCGAACATCAAGCTGCTAGACTTCACTTCTCACTCGAA	323
Db	765	TRYKKMMAMTWMSWRMKSYRMSGMRMMSAMRYCSRKAKCTKYASARMTKRAK	706
QY	324	GTTACGATATATCCGTTTCACACATACGGGAAGTAGAGACTTGAAATTAAGATTAA	383
Db	705	RSYRFRRRMYMKRKMWTYRYRYMRSCMTYRPAWBSRRKAGASMSLMMYMRGASMYWS	646
QY	384	CTCCCTGACGAAGTTAAAGTGTGGGGTACGACGGGATAGGTAAACCGCAATGGTGC	443
Db	645	KVSCBAKCKCKTRYTSSYMSYMGMYSSYSKMSMTSMTSMTGMDTCTMYTSMKSTSR	586
QY	444	GTTTGGGTTATCAATCATCAACATTAAGAAATGACGATCATCCATAGAGAGTGGTTA	503
Db	505	SMGMSWGSMSRWYMMWKKMRGRKMYRYMKCTYMRCHCTRYGTTMYTSSRMNYTR	526
QY	504	TGGCATAT-TCCATTTGCGAAGCTTAAACGGGAAGCTTATCAGCCGGAATTCAGTCTG	562
Db	525	YKARTYSKRYMYKYRKYCMYYGYMKCSYMRVGYCKAKCKCYAMCKAAVSGM	466
QY	563	AGATATATGCAAAATCAAATTCAGGACATATCCACCGAAGCTCAACCTGGTGTGCA	622
Db	465	MTMYKYSMMWMSSTKYMMSMYKYKCRSMKYAKCYGCKMYTYSYGMKMYTMMGSYK	406
QY	623	TGACACAAAATCTCGTATTTGATTCGCGGATTTTATGACATATGGTCAAGTCGGGACAC	682
Db	405	YSRCIKYMRMYMKMMYMYSAISSMMTYYYAAKYMKYKRGTMSSWGSYKXY	346
QY	683	ATTACTCAC	691
Db	345	CTMWCCTMCK	337

RESULT 5
ADA71938
ID ADA71938 standard; DNA; 2000 BP.

AC ADA719387

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

KW Plant; bacterial infection; fungal infection; viral infection; rice; KW gene; ds.

OS *Oryza sativa*.

```
XX XX MO2003000898-A1.
XX PN
XX PD 03-JAN-2003.
XX XX
XX PF 22-JUN-2001; 2001WO-1B001105.
XX PR 22-JUN-2001; 2001WO-1B001105.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX PI Katagiri F, Khan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX DR MPI; 2003-175290/17.
XX PT Identifying at least one gene involved in plant resistance or response to
XX PT pathogenic infection for conferring resistance or tolerance to a plant to
XX PT bacterial, fungal or viral infection by determining or detecting plant
XX PT gene expression.
XX PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX CC The present invention relates to a method (M1) for identifying genes
XX CC involved in plant resistance or response to pathogenic infection. M1
XX CC comprises identifying a gene whose expression is significantly altered in
XX CC the incompatible interaction of plant gene expression relative to
XX CC expression of the gene in an uninfected plant, in a mutant plant that
XX CC does not express a gene associated with response to pathogenic infection,
XX CC or in a corresponding incompatible or compatible interaction. (M1) is
XX CC useful for conferring resistance to resistance or tolerance to a plant to
XX CC bacterial, fungal or viral infection. The present sequence was used to
XX CC illustrate the invention.
XX SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match          4.9%; Score 36.6; DB 8; Length 2000;
Best Local Similarity 8.2%; Pred. NO. 0.73; Mismatches 7; Gaps 3;
Matches 49; Conservative 290; Indels 249; Indels 7; Gaps 3;

QY 82 CTGCGTTTGGCCGTTTCATTTGGGCAATTAATTTCTTCCGCAACAGGTATATGATT 141
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 256 SBRBSMMWMMKMRKRSYSGWYSKYMCTAYKXSYRSKCYRRGGRGRRYRG 315
QY 142 CGAGGAAACACCCGCGAACGATCAACGTTAAGCTTCTT-----TGGTTA 196
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 316 RGTSMAMMYKMYRKYRGKMGKRWMMRMCRMSKXACYYMRMRMRMTBRBR 375
QY 197 ATTGACGAATAGTTCCAGAGTAGCAAGCAAGCAAAATCTGAGTACGTCTCCCGCTC 256
DB   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 376 AKKSSPTSRKRRKRWKMRKYSRMSCKRARRMKCRSGRARRMCGRGMTGR 435
QY 257 ATTAAGCTCTCA-ACCTATTGCAACATGCGCAACATCACTGCTCACTTCTTCT 315
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 436 MTSYGMRRKMSKMRKMSKMSRMRKXKCSRTTMMKTRGGMGMGCRKRYKRG 495
QY 316 CAGTCAAGTAC-GATATATCCGTTTCAACATACGGAAGTAGAGACTTGGAAAT 374
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 496 MKRCKRRRWGMYRMRKYSARVYTRCARKKYSAAARKARCMVRGXYWAGMW 555
QY 375 TAGAGTAACTCCCTGCAAGATTAAAGTGTGGGTACGACGGGATAGTGAACCGC 434
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 556 MRRYRMYKMMWYKRYSKCSWYCKMSYASCKRSKRGAKRCKRSKMSKMSMR 615
QY 435 AAATGTCGTTTGGGTTATCAACATCAACATTAAGAATGACGATCCATATGA 494
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 616 SSRKCKKSKSSAKRYAMMGMSGMSRMSKSTCYMRKMSKSTCTMYMYMSKT 675
QY 495 GGTCTGTTAGCATATTCATTGCAAGATTAAACGGGAACTTTATCACCGGATTC 554
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 676 YAKYGYMYRYRMYCMWYRYRYRSTYMAVYTSSTYMAVYTSSTYMAVYTSSTY 735
QY 555 TACTGTGAGATATATGCAAAATCAAAATTCAGGAGTATTCGACCGACTCAACCTCG 614
```

```
DB 736 YCKCSWKRSMTMYMSWMAKTWKMRKRYATMMWMMYYSKMYTWCTMMGTYMTWK 795
QY 615 TGTGTCATGACGACAAATCTGTATTTGATATGCCGATTTATGACATATGCT 669
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 796 YMKRMYMYKCTKYTWYMGATYTWTGTAAMWMAKTMRMGMTGAKTRGRARARY 850

RESULT 6
ACC47595
ID ACC47595 standard; DNA; 2306 BP.
XX AC ACC47595;
XX DT 23-JUN-2003 (first entry)
XX DE Gastrodia elata antifungal protein GFP-encoding gene.
XX KW Gastrodia antifungal protein; GFP; fungus-inducible promoter; plant;
XX KW transgenic; gene; ds.
XX OS Gastrodia elata.
XX FH Key location/Qualifiers
FH misc_feature 1..1822
FT /*tag= a
FT /note= "This region is specifically claimed in claim 1"
FT promoter 1..1305
FT /*tag= b
FT /note= "Specifically claimed in claim 2"
FT CDS 1307..1822
FT /*tag= c
FT /product= "GFP"
FT /note= "Gastrodia antifungal protein"
XX PN CN1366045-A.
XX PD 28-AUG-2002.
XX PF 16-JAN-2001; 2001CN-00100478.
XX PR 16-JAN-2001; 2001CN-00100478.
XX PA (GENE-) INST GENETICS CHINESE ACAD SCI.
XX PI Sun Y, Sha Q, Li W;
XX PI Sun Y, Sha Q, Li W;
XX DR MPI; 2003-240396/24.
XX PT Inducible promoter of fungus.
XX PS Claim 1; Page 7-8 (Disclosure); 11pp; Chinese.
XX CC The invention relates to a fungus-inducible promoter from the plant
XX CC Gastrodia elata. The promoter is that of the Gastrodia antifungal protein
XX CC (GAP) gene which is expressed in the tubers of the plant. The fungus-
XX CC inducible promoter may be used in transgenic plants to mediate expression
XX CC of heterologous genes in the presence of fungus. The present sequence
XX CC represents the GAP gene sequence which includes its promoter
XX SQ Sequence 2306 BP; 735 A; 416 C; 441 G; 714 T; 0 U; 0 Other;

Query Match          4.9%; Score 36.2; DB 8; Length 2306;
Best Local Similarity 57.5%; Pred. NO. 1.1; Mismatches 48; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 379 GTTAATCTCCCTGCAAGATTAAAGTGTGGGTACGACGGGATGTTGAAACCGCAAT 438
DB   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1631 GGTAACTACTATCTGATCTTCAAGAGATGTAACGTCGCTATTAAGTATTTCTAT 1690
QY 439 GGTGCGTTTGGGTTATCAACATCAACATTAAGAATGACGATTCATCTCTAT 491
DB   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1691 AATGCAATTTGGGCAACCAACGACGATTTGGAATCTGAATATCATCTCAT 1743
```

```
RESULT 7
ADA13411.1
Continuation (2 of 4) of ADA13411 from base 100001 (Human transporter protein genomic DN
WP Sequence split into 4 fragments LOCUS ADA13411 Accession Ada13411
WP Fragment Name Begin End
WP ADA13411_0 1 110000
WP ADA13411_1 100001 210000
WP ADA13411_2 200001 310000
WP ADA13411_3 300001 389627

Query Match
Best Local Similarity 4.9%; Score 36.2; DB 9; Length 110000;
Matches 62; Conservativity 0; Mismatches 43; Indels 0; Gaps 0;

QY 102 GGGGCAATTAATGCTTCGCGAACAAGGTATATGATTGAGGAAACACCGCGGAA 161
DB 56226 GAGGGAACAGATTCTCTTACAAATGCCATGTTGTGTTCTTGAAAAGCAGCAAAA 56285
QY 162 CGCATCAACGTTAGCGCTACAACTTCTTGTTAATTCGACGAA 206
DB 56286 CTTTCAAGGCAACGTTTCAATTCGGAATACCAAGAA 56330

RESULT 8
ADQ59446.1
Continuation (2 of 4) of ADQ59446 from base 100001 (Human cancer-associated (CA) gene se
WP Sequence split into 4 fragments LOCUS ADQ59446 Accession Adq59446
WP Fragment Name Begin End
WP ADQ59446_0 1 110000
WP ADQ59446_1 100001 210000
WP ADQ59446_2 200001 310000
WP ADQ59446_3 300001 405660

Query Match
Best Local Similarity 4.9%; Score 36.2; DB 12; Length 110000;
Matches 62; Conservativity 0; Mismatches 43; Indels 0; Gaps 0;

QY 102 GGGGCAATTAATGCTTCGCGAACAAGGTATATGATTGAGGAAACACCGCGGAA 161
DB 64226 GAGGGAACAGATTCTCTTACAAATGCCATGTTGTGTTCTTGAAAAGCAGCAAAA 64285
QY 162 CGCATCAACGTTAGCGCTACAACTTCTTGTTAATTCGACGAA 206
DB 64286 CTTTCAAGGCAACGTTTCAATTCGGAATACCAAGAA 64330

RESULT 9
ADZ13757.1
Continuation (2 of 4) of ADZ13757 from base 100001 (Human cancer-associated genomic DNA
WP Sequence split into 4 fragments LOCUS ADZ13757 Accession Adz13757
WP Fragment Name Begin End
WP ADZ13757_0 1 110000
WP ADZ13757_1 100001 210000
WP ADZ13757_2 200001 310000
WP ADZ13757_3 300001 408438

Query Match
Best Local Similarity 4.9%; Score 36.2; DB 14; Length 110000;
Matches 62; Conservativity 0; Mismatches 43; Indels 0; Gaps 0;

QY 102 GGGGCAATTAATGCTTCGCGAACAAGGTATATGATTGAGGAAACACCGCGGAA 161
DB 64474 GAGGGAACAGATTCTCTTACAAATGCCATGTTGTGTTCTTGAAAAGCAGCAAAA 64533
QY 162 CGCATCAACGTTAGCGCTACAACTTCTTGTTAATTCGACGAA 206
DB 64534 CTTTCAAGGCAACGTTTCAATTCGGAATACCAAGAA 64578

RESULT 10
ABAI7217
?
```

```
ID ABAI7217 standard; DNA; 7636 BP.
AC
XX
XX ABAI7217;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 9548.
XX
XX Homo sapiens.
OS
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 11-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 14-JUL-2000; 2000US-0217496P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.
XX
XX 14-AUG-2000; 2000US-0225447P.
XX
XX 14-AUG-2000; 2000US-0225757P.
XX
XX 14-AUG-2000; 2000US-0225758P.
XX
XX 14-AUG-2000; 2000US-0225759P.
XX
XX 18-AUG-2000; 2000US-0226279P.
XX
XX 22-AUG-2000; 2000US-0226681P.
XX
XX 22-AUG-2000; 2000US-0226868P.
XX
XX 22-AUG-2000; 2000US-0227183P.
XX
XX 23-AUG-2000; 2000US-0227009P.
XX
XX 30-AUG-2000; 2000US-0228924P.
XX
XX 01-SEP-2000; 2000US-0228928P.
XX
XX 01-SEP-2000; 2000US-0229343P.
XX
XX 01-SEP-2000; 2000US-0229344P.
XX
XX 01-SEP-2000; 2000US-0229345P.
XX
XX 05-SEP-2000; 2000US-0229509P.
XX
XX 05-SEP-2000; 2000US-0229513P.
XX
XX 06-SEP-2000; 2000US-0230438P.
XX
XX 08-SEP-2000; 2000US-0231242P.
XX
XX 08-SEP-2000; 2000US-0231243P.
XX
XX 08-SEP-2000; 2000US-0231244P.
XX
XX 08-SEP-2000; 2000US-0231413P.
XX
XX 08-SEP-2000; 2000US-0231414P.
```


PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234979P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244221P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249307P.
PR 17-NOV-2000; 2000US-0249308P.
PR 17-NOV-2000; 2000US-0249309P.
PR 17-NOV-2000; 2000US-0249310P.
PR 17-NOV-2000; 2000US-0249311P.
PR 17-NOV-2000; 2000US-0249312P.
PR 17-NOV-2000; 2000US-0249313P.
PR 17-NOV-2000; 2000US-0249314P.
PR 17-NOV-2000; 2000US-0249315P.
PR 17-NOV-2000; 2000US-0249316P.
PR 17-NOV-2000; 2000US-0249317P.
PR 17-NOV-2000; 2000US-0249318P.
PR 17-NOV-2000; 2000US-0249344P.
PR 17-NOV-2000; 2000US-0249345P.
PR 17-NOV-2000; 2000US-0249346P.
PR 17-NOV-2000; 2000US-0249365P.
PR 17-NOV-2000; 2000US-0249379P.
PR 17-NOV-2000; 2000US-0249399P.

PR 17-NOV-2000; 2000US-0249309P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Rosen CA, Barash SC, Ruben SM,
PI
XX WPI, 2001-541565/60.
DR
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX
PS Disclosure; SEQ ID NO 9548; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 7636 BP; 2263 A; 1773 C; 1517 G; 2083 T; 0 U; 0 Other;
Query Match 4.8%; Score 35.6; DB 5; Length 7636;
Best Local Similarity 54.6%; Pred. No. 3;
Matches 71; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
QY 209 GTTCACAGTAGCAAGAAGAGCAAACTCAGTAGCTCTCCGCTCATAGCTACGA 268
DB 2992 GTCCACATATGGGAGACAGTAGCAACGACTGTACACTTACACACAGGACATAGCA 3041
QY 269 ACTCATTTGCAACATCGGCAATCATGCTGTACACATTCATCTTCTCAGTGAAGTTAC 328
DB 3042 CTTCCCTCTCCAGCTTAATCCCATATGAATCAGGTCACTTTGGCTCAGATTAT 3101
QY 329 GATATATCCC 338
DB 3102 TATATATCCC 3111
RESULT 11
ADKS6798/C
ID ADKS6798 standard; DNA; 580 BP.
XX
AC ADKS6798;
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Plant DNA sequence which confers altered metabolic characteristic #4181.
DE

Oy	475	GAGGATCCATCTTATGAGGTGTTGATGGCATATTCATTTCAGAACTTAACGGG	534
Db	1330	AAGGATCTTATGCTAATGAGATGATGAAGTGAACGAATGCTCAACTAAACGCA	1443
Oy	535	AAC 537	
Db	1450	AAC 1452	
RESULT 13			
ID	AAK66772		
XX	AAK66772	standard; DNA; 4248 BP.	
AC	AAK66772;	.	
XX			
DT	06-NOV-2001	(first entry)	
XX			
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21584.		
XX			
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
KW	cytotoxic; gene therapy; vaccine; metastasis; ds.		
OS	Homo sapiens.		
XX			
PN	WO200157182-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	17-JAN-2001;	2001WO-US001354.	
XX			
PR	31-JAN-2000;	2000US-0179065P.	
PR	04-FEB-2000;	2000US-0180628P.	
PR	24-FEB-2000;	2000US-0184664P.	
PR	02-MAR-2000;	2000US-0186350P.	
PR	16-MAR-2000;	2000US-0189874P.	
PR	17-MAR-2000;	2000US-0190076P.	
PR	18-APR-2000;	2000US-0198123P.	
PR	19-MAY-2000;	2000US-0205515P.	
PR	07-JUN-2000;	2000US-0209467P.	
PR	28-JUN-2000;	2000US-0214886P.	
PR	30-JUN-2000;	2000US-0215135P.	
PR	07-JUL-2000;	2000US-021647P.	
PR	07-JUL-2000;	2000US-0216880P.	
PR	11-JUL-2000;	2000US-0217487P.	
PR	11-JUL-2000;	2000US-0217496P.	
PR	14-JUL-2000;	2000US-0218290P.	
PR	26-JUL-2000;	2000US-0220963P.	
PR	26-JUL-2000;	2000US-0220964P.	
PR	14-AUG-2000;	2000US-0224518P.	
PR	14-AUG-2000;	2000US-0224519P.	
PR	14-AUG-2000;	2000US-0225213P.	
PR	14-AUG-2000;	2000US-0225214P.	
PR	14-AUG-2000;	2000US-0225266P.	
PR	14-AUG-2000;	2000US-0225267P.	
PR	14-AUG-2000;	2000US-0225268P.	
PR	14-AUG-2000;	2000US-0225270P.	
PR	14-AUG-2000;	2000US-0225447P.	
PR	14-AUG-2000;	2000US-022547P.	
PR	14-AUG-2000;	2000US-0225757P.	
PR	14-AUG-2000;	2000US-0225758P.	
PR	14-AUG-2000;	2000US-0225759P.	
PR	18-AUG-2000;	2000US-0226279P.	
PR	22-AUG-2000;	2000US-0226681P.	
PR	22-AUG-2000;	2000US-0226686P.	
PR	22-AUG-2000;	2000US-0227182P.	
PR	23-AUG-2000;	2000US-0227099P.	
PR	30-AUG-2000;	2000US-0228924P.	
PR	01-SEP-2000;	2000US-0229287P.	
PR	01-SEP-2000;	2000US-0229343P.	
PR	01-SEP-2000;	2000US-0229344P.	
PR	01-SEP-2000;	2000US-0229345P.	
PR	05-SEP-2000;	2000US-0229509P.	
PR	05-SEP-2000;	2000US-0229513P.	
PR	06-SEP-2000;	2000US-0230437P.	

PR	06-SEP-2000	2000US - 0230438P
PR	08-SEP-2000	2000US - 02311242P
PR	08-SEP-2000	2000US - 02311243P
PR	08-SEP-2000	2000US - 02311244P
PR	08-SEP-2000	2000US - 02311413P
PR	08-SEP-2000	2000US - 02311414P
PR	08-SEP-2000	2000US - 0232080P
PR	08-SEP-2000	2000US - 0232081P
PR	12-SEP-2000	2000US - 0231966P
PR	14-SEP-2000	2000US - 0232337P
PR	14-SEP-2000	2000US - 0232339P
PR	14-SEP-2000	2000US - 0232340P
PR	14-SEP-2000	2000US - 0232401P
PR	14-SEP-2000	2000US - 0233063P
PR	14-SEP-2000	2000US - 0233064P
PR	14-SEP-2000	2000US - 0233065P
PR	21-SEP-2000	2000US - 0234223P
PR	21-SEP-2000	2000US - 0234977P
PR	21-SEP-2000	2000US - 0234978P
PR	25-SEP-2000	2000US - 0234998P
PR	26-SEP-2000	2000US - 0235484P
PR	27-SEP-2000	2000US - 0235836P
PR	27-SEP-2000	2000US - 0235837P
PR	29-SEP-2000	2000US - 0236337P
PR	29-SEP-2000	2000US - 0236367P
PR	29-SEP-2000	2000US - 0236368P
PR	29-SEP-2000	2000US - 0236369P
PR	29-SEP-2000	2000US - 0236370P
PR	02-OCT-2000	2000US - 0236802P
PR	02-OCT-2000	2000US - 0237037P
PR	02-OCT-2000	2000US - 0237038P
PR	02-OCT-2000	2000US - 0237039P
PR	02-OCT-2000	2000US - 0237040P
PR	13-OCT-2000	2000US - 0239935P
PR	13-OCT-2000	2000US - 0239937P
PR	20-OCT-2000	2000US - 0240960P
PR	20-OCT-2000	2000US - 0241121P
PR	20-OCT-2000	2000US - 0241122P
PR	20-OCT-2000	2000US - 0241185P
PR	20-OCT-2000	2000US - 0241178P
PR	20-OCT-2000	2000US - 0241808P
PR	20-OCT-2000	2000US - 0241809P
PR	01-NOV-2000	2000US - 0241826P
PR	01-NOV-2000	2000US - 0244647P
PR	08-NOV-2000	2000US - 0246474P
PR	08-NOV-2000	2000US - 0246475P
PR	08-NOV-2000	2000US - 0246476P
PR	08-NOV-2000	2000US - 0246477P
PR	08-NOV-2000	2000US - 0246527P
PR	08-NOV-2000	2000US - 0246528P
PR	08-NOV-2000	2000US - 0246532P
PR	08-NOV-2000	2000US - 0246534P
PR	08-NOV-2000	2000US - 0246556P
PR	08-NOV-2000	2000US - 0246556P
PR	08-NOV-2000	2000US - 0246611P
PR	08-NOV-2000	2000US - 0246613P
PR	17-NOV-2000	2000US - 0249207P
PR	17-NOV-2000	2000US - 0249208P
PR	17-NOV-2000	2000US - 0249209P
PR	17-NOV-2000	2000US - 0249210P
PR	17-NOV-2000	2000US - 0249211P
PR	17-NOV-2000	2000US - 0249212P
PR	17-NOV-2000	2000US - 0249213P
PR	17-NOV-2000	2000US - 0249214P
PR	17-NOV-2000	2000US - 0249215P
PR	17-NOV-2000	2000US - 0249216P
PR	17-NOV-2000	2000US - 0249217P
PR	17-NOV-2000	2000US - 0249218P

Sequence	Accession	Species	Gene	Protein	Function	Notes
ADX12170	standard	CDNA	729 BP			
ADX12170	(first entry)					
Plant full length insert polynucleotide seqid 6745.						
plant protectant; plant growth regulant; gene therapy; plant; recombinant DNA construct; physical array; plant breeding marker; cold tolerance; heat tolerance; drought tolerance; herbicide tolerance; extreme osmotic condition; pathogen tolerance; pest tolerance; growth rate; cell cycle pathway; disease resistance; galactomannan production; lignin production; plant growth regulator; yield; plant growth; plant development; seed oil; protein yield; protein content; gene; ss.						
Identified.						
US2004034888-A1.						
19-FEB-2004.						
28-APR-2003; 2003US-00425114.						
06-MAY-1999; 99US-00304517.						
05-NOV-2001; 2001US-00985678.						
(LNUJ/) LNU J.						
(ZHOU/) ZHOU Y.						
(KOVA/) KOVALIC D K.						
(SCRE/) SCREEN S E.						
(TABAS/) TABASKA J E.						
(CAOY/) CAO Y.						
Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;						
WPI; 2004-180133/17.						
New recombinant DNA construct, useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for conferring increased resistance to plant disease, or for improving yield.						
Claim 1; SEQ ID NO 6745; 15bp; English.						
The invention describes a recombinant DNA construct comprising a polynucleotide consisting of a sequence encoding an amino acid sequence available in electronic form from the US patent office at http://seqdata.uspto.gov/sequence.html?docID:2004034888 . The polynucleotide of the invention are also useful in physical arrays of molecules and as plant breeding markers. The recombinant DNA construct is useful for improving plant tolerance to cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or pests, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for conferring increased resistance to plant disease, for producing galactomannan, lignin or plant growth regulators, for increasing the rate of homologous recombination in plants, for improving yield by modification of photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake or by providing improved plant growth and development under at least one stress condition or for modifying seed oil or protein yield and/or content. This sequence represents a plant full length insert polynucleotide that can be used in the recombinant DNA construct of the invention.						
Sequence 729 BP; 208 A; 152 C; 171 G; 198 T; 0 U; 0 Other;						
Query Match	4.5%	Score 33.8;	DB 13;	Length 729;		
Best Local Similarity	52.5%;	Pred. No. 3.9;				
Matches 74; Conservative 0;	Mismatches 67;	Indels 0;	Gaps 0;			

Qy	584	JAGGAGCATATTCGACCGACCTCAACCTCGTGTGTTCATGACGAATTCGTATTTG	643
Db	464	CTGGGCATACATGCACCGAGTGAACAAGCCACCGAGTGTCCAAACGATTTAGGAATGGG	523
Qy	644	ATATGCGCGATTTTATGACATATGTGTCACTGTGGGACGATTACTGCTTGTCC	703
Db	524	TTTCACGGAGATCTCTCTGATATTAAGTGTCTGAGAGATTACGTACGTGTGCTCTA	583
Qy	704	TGGGTTTCTTCGGGTCAAGT	724
Db	584	TGAGGTAGGCCGACGTACG	604

Search completed: April 7, 2006, 21:33:54
Job time : 597 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 21:26:27 ; Search time 3831 Seconds

(without alignments)
9086.293 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744
Sequence: 1 gtcgcaattatgaagttt.....aaacgacgtatgcttcgcg 744

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lifting first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.8	5.3	444	1	AU269824 AU269824
2	39.6	5.3	714	6	CA996800 rg03d03.y
3	39.2	5.3	475	3	BU445212 BU445212
4	39.2	5.3	729	5	BU364247 BU364247
5	39.2	5.3	782	3	BU436673 BU436673
6	39	5.2	393	3	BU439058 BU439058
7	39	5.2	742	8	DR070622 RTDK1.14
8	39	5.2	782	7	CO167205 FLU1.67.C
9	38.6	5.2	374	3	BU440511 BU440511
10	38.6	5.2	391	3	BU440677 BU440677
11	38.6	5.2	393	3	BU440820 BU440820
12	38.6	5.2	394	3	BU440720 BU440720
13	38.6	5.2	422	3	BU437922 BU437922
14	38.6	5.2	458	3	BU444052 BU444052
15	38.6	5.2	463	3	BU438307 BU438307
16	38.6	5.2	463	3	BU445358 BU445358
17	38.6	5.2	470	3	BU438852 BU438852
18	38.6	5.2	472	3	BU437919 BU437919
19	38.6	5.2	503	3	BU444250 BU444250
20	38.6	5.2	506	3	BU442298 BU442298
21	38.6	5.2	524	3	BU443637 BU443637
22	38.6	5.2	524	3	BU443637 BU443637

c	23	38.6	5.2	526	3	BU441822	BU441822
c	24	38.6	5.2	531	1	AU268364	AU268364
c	25	38.6	5.2	532	3	BU439175	BU439175
c	26	38.6	5.2	535	3	BU442262	BU442262
c	27	38.6	5.2	539	3	BU444947	BU444947
c	28	38.6	5.2	541	3	BU444725	BU444725
c	29	38.6	5.2	543	3	BU446010	BU446010
c	30	38.6	5.2	547	3	BU437840	BU437840
c	31	38.6	5.2	555	3	BU444817	BU444817
c	32	38.6	5.2	557	3	BU442366	BU442366
c	33	38.6	5.2	558	3	BU443146	BU443146
c	34	38.6	5.2	558	3	BU446802	BU446802
c	35	38.6	5.2	567	3	BU437291	BU437291
c	36	38.6	5.2	570	3	BU442058	BU442058
c	37	38.6	5.2	571	3	BU437532	BU437532
c	38	38.6	5.2	572	3	BU442947	BU442947
c	39	38.6	5.2	573	3	BU445870	BU445870
c	40	38.6	5.2	575	3	BU446620	BU446620
c	41	38.6	5.2	578	3	BU436557	BU436557
c	42	38.6	5.2	578	3	BU438536	BU438536
c	43	38.6	5.2	578	3	BU44484	BU44484
c	44	38.6	5.2	579	3	BU442001	BU442001
c	45	38.6	5.2	580	3	BU446126	BU446126

ALIGNMENTS

RESULT 1
AU269824 444 bp mRNA linear EST 26-APR-2004

LOCUS AU269824 VS Dictyostellium discoideum cDNA clone VSJ372 3', mRNA

DEFINITION

sequence.

ACCESSION

AU269824.1

VERSION

GI:20528622

KEYWORDS

EST

SOURCE

Dictyostellium discoideum

ORGANISM

Dictyostellium discoideum

REFERENCE

1 (bases 1 to 444)

AUTHORS

Urushihara,H., Morio,T., Saito,T., Kohara,Y., Koziki,E., Ochiai,H.,

TITLE

Analyses of cDNAs from growth and slug stages of Dictyostellium

JOURNAL

Nucleic Acids Res. 32 (5), 1647-1653 (2004)

COMMENT

15010511

CONTACT

Hideko Urushihara

INSTITUTE

Institute of Biological Sciences

UNIVERSITY

University of Tsukuba

1-1-1 Tennoudai,

Tsukuba, Ibaraki 305-8572, Japan

TEL

Tel: 81-298-53-4664

FAX

Fax: 81-298-53-6614

EMAIL

Email: hideko@biol.tsukuba.ac.jp.

LOCATION

location/Qualifiers

FEATURES

source

1..444

/organism="Dictyostellium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="VSJ372"

/dev_stage="vegetative"

/sex="mat A"

/clone_lib="VS"

ORIGIN

Query Match

5.3%; Score 39.8; DB 1; Length 444;

Best Local Similarity

51.4%; Pred. No. 0.51;

Matches 89; Conservative

0; Mismatches 84; Indels 0; Gaps 0;

422 TAGGTGAACCGCAATGTCGTTTGGTTATCAATCAATCAATGAATGCGAT

481

70 TGGTTGAAGAAAGAAATTAAGTATGATCAACAAACCGTATTAAATGCGAAT

129

QY 482 CCACCTCTAATGAGGTCGTGATGAGCATATTCATTGCAAGATCTTAAACGGAAAGCTTT 541
 DB 130 CTTTGTGTAATGATCNCCAGGTCGTGATGCTTAAATGGGTAATGGTAAAGATG 189
 QY 542 ATCAAGCGGATTCCTGCTGATGATATGCAATATCAATTCAGGACTATT 594
 DB 190 CAATTCATATGTTACTGTGTGAAGTTATGCTATCAATCTCTATAAAT 242

RESULT 2

CA996800/c

LOCUS CA996800 714 bp mRNA linear EST 07-JAN-2003
 DEFINITION rg03d03.y1 Meiodogyne hapla J2 PAMP1 v1 Meiodogyne hapla cDNA 5' similar to TR:001685 001685 SIMILAR TO ALANINE AMINOTRANSFERASE.

ACCESSION CA996800
 VERSION CA996800.1 GI:27541671
 KEYWORDS EST.

ORGANISM Meiodogyne hapla
 SOURCE Meiodogyne hapla

REFERENCE
 AUTHORS

1 (bases 1 to 714)
 McCarter, J., Clifton, S., Chapell, B., Page, D., Martin, J.,
 Wylie, T., Danne, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritzer, B., Bennett, J., Franklin, C.,
 Tsagaris, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterson, R. and Wilson, R.

TITLE The Washington Univ. Nematode EST Project, 1999
 JOURNAL Unpublished (1999)
 COMMENT

The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu

The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. J2 were provided by Dr. Valerie Williamson of the University of California at Davis (vwilliamson@ucdavis.edu).
 Seg primer: -40RP from Gibco
 High quality sequence stop: 414.

FEATURES

source

1..714
 /organism="Meiodogyne hapla"
 /mol_type="mRNA"
 /db_xref="taxon:6305"
 /dev_stage="J2"
 /lab_host="DH10B"
 /clone_lib="Meiodogyne hapla J2 PAMP1 v1"
 /note="Vector: PAMP1 (Gibco); Site 1: NotI; Site 2: SalI;
 The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UD3 sites of PAMP1. J2 were provided by Dr. Valerie Williamson of the University of California at Davis (vwilliamson@ucdavis.edu)."

ORIGIN

Query Match 5.3%; Score 39.6; DB 6; Length 714;
 Best Local Similarity 50.5%; Pred. No. 0.69;
 Matches 96; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
 QY 443 CGTTTGGTATTCACATCACCATAAGAAAGACGACCTCTATGAGAGCTGTTG 502
 DB 268 CATTGGATTAATCGTTCACAGCAAGCAAGCTTCTGCAAAATTTGAACGTGTTG 209
 QY 503 ATGGATATTCATTTGACAGAACTTAAACGGAACGTTATGACCGGATCTACTGCTG 562

DB 208 TTGACATATTTGGTTATATGCAATTAATGAATGTAAGTCTTGTCGCTG 149
 QY 563 AGATATATGCAATATCAATTCAGGACTATTCGACCGACTCAACCTGTGTGCA 622
 DB 148 GAATATTAAGTAAACACAGACAGTCGATTAATGCTGCTCAATCTCAAGCTGTGAATGCA 89
 QY 623 TGACGACAAA 632
 DB 88 TTCAAAAAA 79

RESULT 3

BU445212/c

LOCUS BU445212 475 bp mRNA linear EST 13-MAR-2002
 DEFINITION BU445212 Dictyostelium discoideum cDNA library, VF Dictyostelium discoideum cDNA clone dv58e21 3', mRNA sequence.

ACCESSION BU445212
 VERSION BU445212.1 GI:19419933
 KEYWORDS EST.

ORGANISM Dictyostelium discoideum
 SOURCE Dictyostelium discoideum

REFERENCE
 AUTHORS Eukaryota; Mycetozoa; Dictyostellidae; Dictyostelium.
 1 (bases 1 to 475)
 Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-1, T.

TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
 JOURNAL Unpublished (2002)
 COMMENT

Contact: Tadazu Shin-1
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES

source

1..475
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="dv58e21"
 /sex="mat A"
 /dev_stage="Growth phase"
 /clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 5.3%; Score 39.2; DB 3; Length 475;
 Best Local Similarity 51.4%; Pred. No. 0.81;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 422 TAGGTGAACCGCAATGTCGCTTTGGTTATCAACATCACCTTAAGAAATGCGAAT 481
 DB 269 TGGTGAAGAAAGAAATTAAGTTATGATATCAACAAACCTATTAAATGCGAAT 210
 QY 482 CCACCTCTATGAGAGTCGTGATGAGCATATTCATTGCAAGACTTAAACGGAAGCTTT 541
 DB 209 CTTTGTTAATGATACCAAGTCGTTTACCTTAATTAATGGGTTTGGTAAAGATG 150
 QY 542 ATCAACCGGATTTCTACTGCTGATATATGCAATATCAATTCAGGACTATT 594
 DB 149 CAATTCATATGTTTACTGTGTAAGTTATGCTCATTAAGTCGCTAAAT 97

RESULT 4

BU364247/c

LOCUS BU364247 729 bp mRNA linear EST 28-NOV-2002
 DEFINITION 603785262F1 CSEQH72 Gallus gallus cDNA clone ChEST740k15 5', mRNA sequence.

ACCESSION BU364247
 VERSION BU364247.1 GI:25872248
 KEYWORDS EST.

SOURCE Gallus gallus (chicken)


```

/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"
ORIGIN

```

Query Match	5.2%	Score 39	DB 3	Length 393
Best Local Similarity	50.0%	Pred. NO.0.88		
Matches 84, Conservative	0	Mismatches 84	Indels 0	Gaps 0

[illegible]

RESULT 7	LOCUS	DEFINITION
DR070622		
DR070622	742 bp	mRNA linear EST 08-JUN-2005
	R1DK1_14_C03.g1_A029	Roots, dark pinus taeda cDNA clone
	R1DK1_14_C03_A029_5'	mRNA sequence.

FEATURES	Location/Qualifiers
source	1. .742

nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24h) prior to harvesting roots for mRNA preparation, the potted trees were placed in a dark growth chamber at 25 C. Double-stranded cDNA was cloned unidirectionally into pSL180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

Query Match	5.24; Score 39; DB 8; Length 742;
Post Total	5.08; Score 39; DB 8; Length 742;

QY	398	TTAAGAGTGTGGGGTACACACGGAAATAGGTGAAACCGCAAAATGTCGCTTTGGGTATCA	457
Db	138	TTAAAGTGAGGAAATGCTCGGAAAGGAAAGTTCTTTTTTTGTGCTTTCAAGATGCA	197
QY	458	ACATCAACATTAAGAAATGACGGATTCATCTCTATGAGAGTGTGATGGCATATTTCCAT	516
Db	198	AACAGGAGAGGTGACATATTCCTCCCAACATTAAGAGAGCGCGTTAAGGCGAGCTTCCAT	256

RESULT 8					
CO167205					
LOCUS	CO167205	782 bp	mRNA	linear	EST 18-JUN-2004
DEFINITION	FLJ1_67_C04.g1 A029 Root flooded Pinus taeda cDNA clone				
	FLJ1_67_C04_A029 5', mRNA sequence.				

FEATURES

```

/organism="Pinus taeda"
/mol_type="mRNA"
/strain="3 CCLONES"
/db_xref="taxon:3352"
/clone="FLD1_67_C04_A029"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Root flooded"
/note="Organ: root; Vector: pSL180; Site_1: EcoRI,
Site_2: XhoI; The library was prepared from polyA+ RNA
from the roots of 1-year-old loblobilly pine (Pinus taeda)
cuttings that were rooted and then planted in washed sand
prior to harvesting tissues for RNA isolation; the rooted
cuttings were maintained for 27 days (April 2003) under

```

ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain 5% soil moisture content. Pots holding the rooted cuttings were fully submerged in water for 24 hours prior to harvest of the roots for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pBl1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 5.2%; Score 39; DB 7; Length 782;
Best Local Similarity 58.0%; Pred. No. 1.1;

Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 398 TTAAAGTGTGGGTCAGACGGAGATAGTGAACCGCAATAGTGGCTTTGGTTATCA 457
DB 284 TTTAAAGTGGGATACGTCCGGAAAGGAGATTCTTTTGTGCTTTCAAGATGCA 343
QY 458 ACATCACCTAAGAAATGAACGATCCATCTCATGAGATGCTGTGATGCGATATTCAT 516
DB 344 AACAGGAAAGTGACATGTCTGCCAACATTAAAGAGGCCGTTAAGGCGACGCTCCAT 402

RESULT 9
LOCUS BJ440426/c 371 bp mRNA linear EST 13-MAR-2002

DEFINITION BJ440426 Dictyostelium discoideum cDNA library, VF Dictyostelium
ACCESSION BJ440426
VERSION BJ440426
KEYWORDS EST. GI:19415148

SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

REFERENCE 1 (bases 1 to 371)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage Unpublished (2002)

JOURNAL Contact: Tadao Shin-i
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES

source

1..371
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv43116"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 371;
Best Local Similarity 51.4%; Pred. No. 1.2;

Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCCGAATGTGCGTTTGGTTATCAACATCAACATTAAGAAATGACGAT 481
DB 252 TGGTGAATAAAGAAATTAAGTTAATGATATCAACAAACCGATTAAGATGCAAT 193
QY 482 CCACTCTATGAGAGTGTGATGCGATATTCATTTCGAGAACTTAAACGGACGTT 541
DB 192 CTTTGTATATATCAACCAAGGTGTTTGTACCTTAATAATGGGATATGGTAAAGATG 133
QY 542 ATCAGCGGATTTACTGCTGAGATATATGAATCAAAATTCAGGACTATT 594

DB 132 CAACCTCAATGTTTACTGCTGAAGTTTATGCTCATTTCAAAATGCTGTAATAAT 80

RESULT 10
LOCUS BJ440571/c 374 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ440571 Dictyostelium discoideum cDNA library, VF Dictyostelium
ACCESSION BJ440571
VERSION BJ440571
KEYWORDS EST. GI:19415293

SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum

REFERENCE 1 (bases 1 to 374)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage Unpublished (2002)

JOURNAL Contact: Tadao Shin-i
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES
source
1..374
Location/Qualifiers

/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv44105"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 374;
Best Local Similarity 51.4%; Pred. No. 1.2;
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCCGAATGTGCGTTTGGTTATCAACATCAACATTAAGAAATGACGAT 481
DB 256 TGGTGAATAAAGAAATTAAGTTAATGATATCAACAAACCGATTAAGATGCAAT 197
QY 482 CCACTCTATGAGAGTGTGATGCGATATTCATTTCGAGAACTTAAACGGACGTT 541
DB 196 CTTTGTATATATCAACCAAGGTGTTTGTACCTTAATAATGGGATATGGTAAAGATG 137
QY 542 ATCAGCGGATTTACTGCTGAGATATATGAATCAAAATTCAGGACTATT 594
DB 136 CAACCTCAATGTTTACTGCTGAAGTTTATGCTCATTTCAAAATGCTGTAATAAT 84

RESULT 11

BJ440677/c

LOCUS BJ440677 391 bp mRNA linear EST 13-MAR-2002
DEFINITION BJ440677 Dictyostelium discoideum cDNA library, VF Dictyostelium
ACCESSION BJ440677
VERSION BJ440677
KEYWORDS EST. GI:19415399

SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum

REFERENCE 1 (bases 1 to 391)
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage Unpublished (2002)

JOURNAL Contact: Tadao Shin-i
COMMENT Center For Genetic Resource Information

RECESSION	08107022	
VERSION	BJ437922.1	GI:19412644
KEYWORDS	EST	

SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 422)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..422
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv35e19"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 422;
Best Local Similarity 51.4%; Pred. No. 1.2;
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGTGAAAACCGCAATGCGTTGGTTATCAACATCAACATTAAGAAATGACGAT 481
DB 247 TGGTGAAGAAAAGAAATATAAAGTTATATATATCAACAAACCGTATTAGATGCGAAT 188
QY 482 CCACCTCTATGAGAGTGGTTGATGCAATTCATTCGAACTTAACGGGAACGTTT 541
DB 187 CTTTGTATATATATCAACCAAGTGGTTTACCTTACATTAATAATGGTATTGGTAAGATG 128
QY 542 ATCAGCCGAGTTCTACTGCTGAGATATATGCAAAATACAAATTCAGGACTATT 594
DB 127 CAATTCAATGTTTACTGTTGAAAGTTATGCTCATTCAAATGCTGTAATAAT 75

RESULT 15
BU444052/c 458 bp mRNA linear EST 13-MAR-2002
LOCUS BU444052 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictoideum cDNA clone ddv55f06 3', mRNA sequence.
ACCESSION BU444052
VERSION BU444052.1 GI:19418773
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 458)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadao Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1..458
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"

/clone="ddv55f06"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 458;
Best Local Similarity 51.4%; Pred. No. 1.2;
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGTGAAAACCGCAATGCGTTGGTTATCAACATCAACATTAAGAAATGACGAT 481
DB 254 TGGTGAAGAAAAGAAATATAAAGTTATATATCAACAAACCGTATTAGATGCGAAT 195
QY 482 CCACCTCTATGAGAGTGGTTGATGCAATTCATTCGAACTTAACGGGAACGTTT 541
DB 194 CTTTGTATATATATCAACCAAGTGGTTTACCTTACATTAATAATGGGTAATGTAAGAATG 135
QY 542 ATCAGCCGAGTTCTACTGCTGAGATATATGCAAAATACAAATTCAGGACTATT 594
DB 134 CAATTCAATGTTTACTGTTGAAAGTTATGCTCATTCAAATGCTGTAATAAT 82

Search completed: April 7, 2006, 22:37:13
Job time : 3836 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 21:31:52 ; Search time 1537 Seconds

(without alignments)
860.446 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744
Sequence: 1 gtgcgaatcagaaagttt.....aaagacgtagctctccg 744

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/PCUTS_COMB.seq:*
 - 7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
 - 8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
 - 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	33.8	4.5	601	US-09-949-016-150404
2	33.8	4.5	181251	Sequence 15970, A
3	33.4	4.5	421118	Sequence 16297, A
4	32.4	4.4	285478	Sequence 13362, A
5	32.2	4.3	1141	Sequence 22, Appl
6	32	4.3	832	Sequence 2813, Ap
7	32	4.3	99304	Sequence 15440, A
8	31.4	4.2	17000	Sequence 18, Appl
9	31.4	4.2	236964	Sequence 15753, A
10	31.2	4.2	8439	Sequence 473, Appl
11	30.8	4.1	1237	Sequence 26, Appl
12	30.8	4.1	1237	Sequence 26, Appl
13	30.8	4.1	1995	Sequence 2645, Ap
14	30.8	4.1	2169	Sequence 3, Appl
15	30.8	4.1	3356	Sequence 3920, Ap
16	30.6	4.1	3001	Sequence 185, App
17	30.4	4.1	601	Sequence 15347, A
18	30.4	4.1	601	Sequence 15347, A
19	30.4	4.1	3509	Sequence 4254, Ap
20	30.4	4.1	3509	Sequence 4255, Ap
21	30.4	4.1	118382	Sequence 15996, A
22	30.4	4.1	118382	Sequence 15997, A
23	30.2	4.1	1194	Sequence 558, Ap
24	30.2	4.1	1194	Sequence 20880, A

C 25	30	4.0	601	3	US-09-949-016-128577	Sequence 128577,
C 26	30	4.0	601	3	US-09-949-016-128578	Sequence 128578,
C 27	30	4.0	601	3	US-09-949-016-150405	Sequence 150405,
C 28	30	4.0	1270	3	US-09-415-277C-1	Sequence 1, Appl
C 29	30	4.0	1270	3	US-09-415-277C-16	Sequence 16, Appl
C 30	30	4.0	1271	3	US-09-415-277C-3	Sequence 3, Appl
C 31	30	4.0	1522	3	US-09-533-559-5608	Sequence 5608, Ap
C 32	30	4.0	1641	2	US-09-080-897-5	Sequence 5, Appl
C 33	30	4.0	1641	3	US-09-323-735-5	Sequence 5, Appl
C 34	30	4.0	8019	3	US-09-949-016-2835	Sequence 2835, Ap
C 35	30	4.0	8019	3	US-09-949-016-2836	Sequence 2836, Ap
C 36	30	4.0	24205	3	US-09-949-016-15385	Sequence 15385, A
C 37	29.8	4.0	351	3	US-09-248-796A-2992	Sequence 2992, Ap
C 38	29.8	4.0	947	3	US-09-640-211A-138	Sequence 138, Ap
C 39	29.8	4.0	1164	3	US-09-248-796A-4125	Sequence 4125, Ap
C 40	29.8	4.0	1443	3	US-09-248-796A-668	Sequence 668, Ap
C 41	29.8	4.0	2090	3	US-09-270-767-4280	Sequence 4280, Ap
C 42	29.8	4.0	2090	3	US-09-270-767-19562	Sequence 19562, A
C 43	29.8	4.0	5619	3	US-09-799-451-241	Sequence 241, Ap
C 44	29.8	4.0	319608	3	US-09-539-333D-1	Sequence 1, Appl
C 45	29.6	4.0	601	3	US-09-949-016-102623	Sequence 102623,

ALIGNMENTS

```

RESULT 1
US-09-949-016-150404
; Sequence 150404, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150404
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-150404

Query Match          4.5%; Score 33.8; DB 3; Length 601;
Best Local Similarity 52.5%; Pred. No. 0.58;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Cy 186 TCTTTGGTAAATTCAGCAATAGTTCACAGTACGAAAGCAAACTCGAGTAC 245
      |||||
Db 131 TTTTATTTTGAAGTGGTCTATTTTAAATTTGAATTAAGCATTAACATTAATTTACTAT 190

Cy 246 GTCCTCCGCTCATAGTCTACGAACTCATTCGACATCGCAATCAAGTCTACAC 305
      |||||
Db 191 TTTAACAATTTTAAAGTGTACAGTACGATTAACAATTAACATTAATGTCAGCTTAC 250

Cy 306 TTCAATCTTCAGTCAAGT 326
      |||||
Db 251 TACCATCATCTCTAAGAACTT 271

RESULT 2
US-09-949-016-15970
; Sequence 15970, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15970
LENGTH: 181251
TYPE: DNA
ORGANISM: Human
US-09-949-016-15970

Query Match 4.5%; Score 33.8; DB 3; Length 181251;

Best Local Similarity 52.5%; Pred. No. 13;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 186 TCTTGTGTTATTCGACGATAGTTCACAGTAGCAAGCAAGCAAACTCGAGTAC 245
Db 58690 TTTTCTTTGAGTGGTCTATTTTAAATGTAATAGCATAAATTAATTTACTAT 58749
Qy 246 GTCTCCGCTCATAGTCTACGAAGCTTCGAAATCGCAATCAAGTGTCTACGAC 305
Db 58750 TTTTAAATTTTAAAGTGTACAGTAGTACAAATTTACATCATGTGACGCTGTAC 58809
Qy 306 TTCACTCTTCAGTCGGAAGT 326
Db 58810 TACCATCCATCTTGAACCTT 58830

RESULT 3

US-09-949-016-16297/C
Sequence 16297, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16297
LENGTH: 421118
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)_(421118)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16297

Query Match 4.5%; Score 33.4; DB 3; Length 421118;

Best Local Similarity 49.2%; Pred. No. 28;
Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 504 TGGCATATTCATTTGCAAGAACTTAACGGAGCACTTTATACCGGATTTACTGCTGA 563
Db 264214 TAGAATCATCATTTGTAAACAAATACGGCATTTTCACAAAAGATTGACTTTTCA 264155

Qy 564 GATATGCAAAATCAATTCAGGACTATTCGACCGACCTGACCTGGTGGTCAT 623
Db 264154 GGTATTTTCTATTATTCGCTAGTATTAAGAACTTAACTTAACTTCTGGGATTA 264095
Qy 624 GACGACAAATCTCGTATTTGATATGCGATTTTATGACATATGTCACGTGGGACG 682
Db 264094 CATTCACGAGAGGCTTTGAATTTTGAATTTTTCACACATGAGATTTATGTTAGC 264036

RESULT 4

US-09-949-016-13362
Sequence 13362, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13362
LENGTH: 285478
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)...(285478)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13362

Query Match 4.4%; Score 32.4; DB 3; Length 285478;

Best Local Similarity 58.2%; Pred. No. 50;
Matches 57; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 249 TCCGCTCATAGTCTGCAAGCTCATTCGAACATGCGCAATCATGCTGCTAGACTTC 308
Db 35275 TCTCCCTCTCAGGCTCTTAAGTATCTGAGACGTGCGCACACACCCGGCTAAGTTT 35334
Qy 309 ATCTTTCAGTGAAGTTACATATATCCCGTTTCACA 346
Db 35335 GTATTTTCGAGAGAGAGGGGTTTCAACATGTTGCCCA 35372

RESULT 5

US-09-806-708B-22/C
Sequence 22, Application US/09806708B
Patent No. 6784342

GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence

FEATURES:
NAME/KEY: promoter
LOCATION: (1)...(1141)

;; TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 6 EXPRESSION
;; FILE REFERENCE: RTS-0187
;; CURRENT APPLICATION NUMBER: US/09/679,299A
;; CURRENT FILING DATE: 2000-10-04
;; NUMBER OF SEQ ID NOS: 164
;; SEQ ID NO 18
;; LENGTH: 17000
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-679-299A-18

Query Match 4.2%; Score 31.4; DB 3; Length 17000;
Best Local Similarity 51.0%; Pred. No. 24;
Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

OY 441 TGGCTTTGGTTATCAACATCACCATTAAGAAATGAGCGATCCCTCTATGAGGTGCT 500
DB 16986 TGCCTTTCTTTTCCATGACAGGCTTAAGAAATCCCTGAACAGCGCGCTCATTTCTTC 16927
OY 501 TGATGCGATATTCATTTGACAGAACTTAACGGGACGTTATGACCGGATTTCTACTGC 560
DB 16926 TGTTCGAATGCAAGTGAAGAACTCAATGAAAAGATTATCTTACAGTTTAAATGT 16867
OY 561 TGAATATATGCAAAATACAAATTC 585
DB 16866 CGTCAGATTTTCATATGATTGA 16842

RESULT 9
US-09-949-016-15753/C
; Sequence 15753, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15753
; LENGTH: 236964
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(236964)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15753

Query Match 4.2%; Score 31.4; DB 3; Length 236964;
Best Local Similarity 50.3%; Pred. No. 98;
Matches 77; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 451 GTTATCAATCATCAGCAATGAAGATCCACTCTATGAGAGTGGTATGAGCATTA 510
DB 4225 GTTATTAATCTTGAATCATAAAGAAAATTAATAATGATGTGTATGTTCTATC 4166
OY 511 TTTCATTTGCAAACTTAACGGGACGTTATATCAGCCGATTTCTACTGCTGAGATATAT 570
DB 4165 TTTCATCTACAGTTGAGAAACGCAATTTTCTAAGCTTAGTATTAACAGTGAAGTTAT 4106
OY 571 GCAATTAACAATTCAGGAGCTATTTCCGACGAC 603
DB 4105 AACCAACATCTAGTTCACAAAAATATTTATGAGC 4073

RESULT 10
US-09-221-017B-473
; Sequence 473, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H.
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 473:
SEQUENCE CHARACTERISTICS:
LENGTH: 8439 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...8439
US-09-221-017B-473

Query Match 4.2%; Score 31.2; DB 3; Length 8439;
Best Local Similarity 51.4%; Pred. No. 19;
Matches 72; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

OY 378 AGTAACTCCCTGACAGAAAGTTAAGAGTGGGTACGACGGATTAAGTGAACCGCAAA 437
DB 5186 AGACAAATTCATTCAGGAATTCAGACAGCGCGCTTAATGAAAACGGTGAACGAAT 5245
OY 438 TGGTGGCTTTGGTTATCAACATCACATAAGAAATGACGAGATCCACTCTATGAGAGT 497
DB 5246 GGGATTTGTGACAGTAAATCAAAAAATATCTCGGCTCGGATGCAATTCAGGCAAAAAGC 5305

QY 498 CGTGAATGCAATTCATT 517
DB 5306 CTATGCCGCAATTTGCTACT 5325

RESULT 11

US-08-808-793-26/c
Sequence 26, Application US/08808793
Patent No. 5858713
GENERAL INFORMATION:
APPLICANT: Soderlund, David M.
APPLICANT: Ingles, Patricia J.
TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
TITLE OF INVENTION: AND USE THEREOF
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSES: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,793
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/034,361
FILING DATE: 24-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,649
FILING DATE: 01-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Braman, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1237 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-808-793-26

Query Match 4.1%; Score 30.8; DB 2; Length 1237;
Best Local Similarity 52.3%; Pred. No. 9.2;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 3 GCGAATTATGAAGTTTGGATTTGGTACCGTATATCATGCTTTATTACAGTG 62
DB 862 GCGTAAAGCGCAATTTTAAATAGCACTTGAATCTGAATATATCATGTT 803
QY 63 GGGGCGAATGACAGCAATTTGCGTTTGGCCGTTTCATTTGGGCAATTAATTTCTTGC 122
DB 802 GAGGTAGTCTGAGACAGTGTGAAGTCTCCGACTGTGTGTATGAGAGGTATCATGT 743
QY 123 GAAACGCGTG 132
DB 742 CAACATGTTG 733

RESULT 12

US-09-710-279-1257/c
Sequence 1257, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1257
LENGTH: 1995
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1257

Query Match 4.1%; Score 30.8; DB 3; Length 1995;
Best Local Similarity 52.3%; Pred. No. 12;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 273 ATTGCAATGCGCAATCAAGCTGCTACAGCTTCATCTTCTCAGTCGAACTTACGATA 332
DB 263 ATTGCAAGCTGCTGATTAATCCCAAGAACTTAGTTTATCATCTGTTTAAAGCG 204
QY 333 TATCCCGTTTCAACATACGGGAAGTAGAGACTTGAATTAAGTTAACTCCCTGCA 392
DB 203 TATCTTATTAACCTTACTTGGCGACATATTAAGCAATTTGAAGTAATGATTTT 144
QY 393 GCAAGTTAAG 402
DB 143 GCGTATTGAG 134

RESULT 13
US-09-134-001C-2645/c
Sequence 2645, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2645
LENGTH: 2169
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2645

Query Match 4.1%; Score 30.8; DB 3; Length 2169;
Best Local Similarity 52.3%; Pred. No. 12;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 273 ATTGCAATGCGCAATCAAGCTGCTACAGCTTCATCTTCTCAGTCGAACTTACGATA 332
DB 437 ATTGCAAGCTGCTGATTAATCCCAAGAACTTAGTTTATCATCTGTTTAAAGCG 378
QY 333 TATCCCGTTTCAACATACGGGAAGTAGAGACTTGAATTAAGTTAACTCCCTGCA 392
DB 377 TATCTTATTAACCTTACTTGGCGACATATTAAGCAATTTGAAGTAATGATTTT 318

QY 393 GCAAGTTAAG 402
|||
Db 317 GCGATTTCAG 308
^

RESULT 14

US-07-998-289B-3/c
; Sequence 3, Application US/07998289B
; Patent No. 6027876
; GENERAL INFORMATION:
; APPLICANT: Black, Bruce C
; APPLICANT: Taylor, Martin
; APPLICANT: Heckel, David G
; TITLE OF INVENTION: Method for Monitoring Pesticide
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESS: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998,289B
; FILING DATE: 30-DEC-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Robinson, Joseph R
; REGISTRATION NUMBER: 33,448
; REFERENCE/DOCKET NUMBER: 0646/OA939
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-998-289B-3

Query Match 4.1%; Score 30.8; DB 3; Length 2279;
Best Local Similarity 52.3%; Pred. No. 13;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 3 GCGAATTATGAAGTTTGGATGATTTGGTACCGTATATCATGCTGTTTATTCAGTG 62
|||
Db 1462 GCGTAAAGGGAACATTTTAAATAGGACCTGAACGATATATCATGATATCATGTT 1403
|||
QY 63 GGGGCGAATGAACAAATTTCTGCGTTTGGCGGTTCAATTTGGGCAATTAATGCTTGGC 122
|||
Db 1402 GAGGTAGTCGAGACAGAGCTGAGAGTCTCCAGCTGCTGATGATGATGAGCGTCATCGT 1343
|||
QY 123 GAAACGCGTG 132
|||
Db 1342 CAACATGTTG 1333
|||

RESULT 15
US-09-710-279-3920/C
; Sequence 3920, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3920
; LENGTH: 3356
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-09-710-279-3920

Query Match 4.1%; Score 30.8; DB 3; Length 3356;
Best Local Similarity 52.3%; Pred. No. 16;
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 273 ATTGCAACATGCGAATCAAGCTGCTAGCACTTCACTTCTGAGTGAAGTACGATA 332
|||
Db 666 ATTGCACTGCTGATTAATTCCTCAAGAACTTAGTTATCACTTCTTTTAAAGCG 607
|||
QY 333 TATCCGTTTCAACATACGGAAGTAGAGACTTGAATTAAGTTAACTCCCTGCA 392
|||
Db 606 TATCTTATTAACCTTACTTGGCGACATATTACTAGCAAAATGGAAGTGAATGATTTT 547
|||
QY 393 GCAAGTTAAG 402
|||
Db 546 GCGATTTCAG 537
|||

Search completed: April 8, 2006, 01:20:33
Job time : 1540 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 22:13:32 ; Search time 823 Seconds
(without alignments)
7475.598 Million cell updates/sec

Title: US-10-784-592-18
Perfect score: 744
Sequence: 1 gtgcgaattatgaagttc.....aaacagcgtatccttcgcg 744

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

Published Applications NA Main: *
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	US-10-784-592-18	Sequence 18, Appl
2	36.2	4.9	389627	US-10-504-689-3	Sequence 3, Appl
3	36.2	4.9	405660	US-10-322-696-82	Sequence 82, Appl
4	35.2	4.7	580	US-10-487-901-4181	Sequence 4181, App
5	35	4.7	2349	US-10-369-963-37301	Sequence 37301, App
6	34.4	4.6	4818	US-10-437-963-63024	Sequence 63024, A
7	34.2	4.6	1502	US-09-925-065A-724617	Sequence 724617,
8	33.8	4.5	660	US-09-925-065A-867215	Sequence 867215,
9	33.8	4.5	591	US-10-425-115-125681	Sequence 125681,
10	33.8	4.5	729	US-10-425-114-6745	Sequence 6745, App
11	33.8	4.5	860	US-10-425-114-6834	Sequence 6834, App
12	33.8	4.5	935	US-10-425-115-154424	Sequence 154424, App
13	33.8	4.5	1240	US-09-925-065A-41609	Sequence 41609, A
14	33.4	4.5	706	US-10-487-901-3702	Sequence 3702, App
15	33.4	4.5	1048	US-10-767-701-11172	Sequence 11172, A
16	33.4	4.5	13077	US-10-915-740A-281	Sequence 281, App
17	33.4	4.5	2242716	US-10-915-740A-1068	Sequence 1068, App
18	33.2	4.5	534	US-09-925-065A-427028	Sequence 427028,
19	33.2	4.5	640	US-09-925-065A-400132	Sequence 400132,
20	33	4.4	1184	US-10-123-155-412	Sequence 412, App
21	33	4.4	1184	US-10-146-731-412	Sequence 412, App
22	33	4.4	1184	US-10-140-472-412	Sequence 412, App
23	33	4.4	1184	US-10-141-761-412	Sequence 412, App

24	33	4.4	1184	6	US-10-142-885-412	Sequence 412, App
25	33	4.4	1184	6	US-10-158-790-412	Sequence 412, App
26	33	4.4	1184	6	US-10-137-871-412	Sequence 412, App
27	33	4.4	1184	6	US-10-140-923-412	Sequence 412, App
28	33	4.4	1184	6	US-10-141-756-412	Sequence 412, App
29	33	4.4	1184	6	US-10-141-759-412	Sequence 412, App
30	33	4.4	1184	6	US-10-140-805-412	Sequence 412, App
31	33	4.4	1184	6	US-10-140-864-412	Sequence 412, App
32	32.8	4.4	641	4	US-09-925-065A-178736	Sequence 178736, App
33	32.6	4.4	566	4	US-09-925-065A-604703	Sequence 604703, Sequence 604704,
34	32.6	4.4	566	4	US-09-925-065A-604704	Sequence 604704,
35	32.6	4.4	2700	7	US-10-683-516-4	Sequence 4, Appl
36	32.4	4.4	611	4	US-09-925-065A-760902	Sequence 760902,
37	32.4	4.4	813	7	US-10-282-122A-16587	Sequence 16587, A
38	32.4	4.4	301477	7	US-10-322-281-456	Sequence 456, App
39	32.2	4.3	563	4	US-09-925-065A-192056	Sequence 192056,
40	32.2	4.3	598	4	US-09-925-065A-301137	Sequence 301137,
41	32.2	4.3	123526	3	US-09-910-185-11	Sequence 11, Appl
42	32	4.3	633	4	US-09-925-065A-914168	Sequence 914168,
43	32	4.3	1470	7	US-10-437-963-88649	Sequence 88649, A
44	32	4.3	26147	10	US-11-097-143-14227	Sequence 14227, A
45	31.8	4.3	503	8	US-10-363-345A-37281	Sequence 37281, A

ALIGNMENTS

RESULT 1
US-10-784-592-18
Sequence 18, Application US/10784592
Publication No. US20050147983A1
GENERAL INFORMATION:
APPLICANT: Willing, Reinhard
APPLICANT: Oestergaard, Peter
APPLICANT: Laessen, Soren
TITLE OF INVENTION: POLYPEPTIDES OF ALICYCLOBACILLUS SP.
FILE REFERENCE: 10406.203-US
CURRENT APPLICATION NUMBER: US/10/784,592
CURRENT FILING DATE: 2004-02-23
NUMBER OF SEQ ID NOS: 55
SOFTWARE: PatentIn version 3.2
SEQ ID NO 18
LENGTH: 744
TYPE: DNA
ORGANISM: Alicyclobacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)_(744)
OTHER INFORMATION: CDS
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(123)
OTHER INFORMATION: sig_peptide
FEATURE:
NAME/KEY: misc feature
LOCATION: (124)..(744)
OTHER INFORMATION: mat_peptide
US-10-784-592-18
Query Match 100.0%; Score 744; DB 9; Length 744;
Best Local Similarity 100.0%; Pred. No. 7.1e-240; Indels 0; Gaps 0;
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGGGAATTATGAAGTTTGGATGATTTGGTACCGTATATCATGCTTTATTCAG 60
DB 1 GTGGGAATTATGAAGTTTGGATGATTTGGTACCGTATATCATGCTTTATTCAG 60
QY 61 TGGGGGGAATGAACAGAAATTCGCTTTTCCGCTTCATTTGGGCATTAAATGCTTC 120
DB 61 TGGGGGGAATGAACAGAAATTCGCTTTTCCGCTTCATTTGGGCATTAAATGCTTC 120
QY 121 GCGAACGCGGTATGATTCGAGAAACACCGCGGAACGATCAACGTAAGCGCT 180
DB 121 GCGAACGCGGTATGATTCGAGAAACACCGCGGAACGATCAACGTAAGCGCT 180

Db	121	CGAACAACGGGTATATGATATTCGAGAAAAACACACCGCGGAAACGCATCAAGTTAACGCT	180
QY	181	ACAACTTCTTGTGTTAATTGACGCAATAGTTTCAAGGTAGCAAAAGCAAGCAAAACTCG	240
Db	181	ACAACTTCTTGTGTTAATTGACGCAATAGTTTCAAGGTAGCAAAAGCAAGCAAAACTCG	240
QY	241	AGTAGCTCCGCGCTCAATAGTCAAGAACATCTTGAATATGCGCAACATCAAGCTGCT	300
Db	241	AGTAGCTCCGCGCTCAATAGTCAAGAACATCTTGAATATGCGCAACATCAAGCTGCT	300
QY	301	ACGACTTCATCTTCTCAGTCGAAAGTTACGATATATCCGTTTCAACATACGCGGAAGSTA	360
Db	301	ACGACTTCATCTTCTCAGTCGAAAGTTACGATATATCCGTTTCAACATACGCGGAAGSTA	360
QY	361	CGAAGCTTGGAAATTAAGATTAACTCCCTCGACGAATTAAAGTGTGGGTACGACGG	420
Db	361	CGAAGCTTGGAAATTAAGATTAACTCCCTCGACGAATTAAAGTGTGGGTACGACGG	420
QY	421	ATAGGTGAAACCGGAATAGTGCGTTTGGGTATCAACATCAACCAATTAAGATGACGA	480
Db	421	ATAGGTGAAACCGGAATAGTGCGTTTGGGTATCAACATCAACCAATTAAGATGACGA	480
QY	481	TCCACTCCTATGAGAGTGTGATGAGCATATTCATTTGAGAACTTAAACGGGAAAGTT	540
Db	481	TCCACTCCTATGAGAGTGTGATGAGCATATTCATTTGAGAACTTAAACGGGAAAGTT	540
QY	541	TATCAGCCCGGAATTTCTAGTCTGAGATATATGCCAAATTAACAAATTCAGGGACTATTTCCGACC	600
Db	541	TATCAGCCCGGAATTTCTAGTCTGAGATATATGCCAAATTAACAAATTCAGGGACTATTTCCGACC	600
QY	601	GACCTCAACCCCTGGGTGTGCCATGACGACCAAAATCGATATTGATATAGCGGAAATTTATG	660
Db	601	GACCTCAACCCCTGGGTGTGCCATGACGACCAAAATCGATATTGATATAGCGGAAATTTATG	660
QY	661	ACATATGTGTCAAGTCGCGGACAGATTACTCACTTGTGCTTCACATGAGTTTCTTCGAGTCA	720
Db	661	ACATATGTGTCAAGTCGCGGACAGATTACTCACTTGTGCTTCACATGAGTTTCTTCGAGTCA	720
QY	721	GATGAAACGACGTATAGTCTTTCCG	744
Db	721	GATGAAACGACGTATAGTCTTTCCG	744

```

RESULT 2
US-10-504-689-3
; Sequence 3, Application US/10504689
; Publication No. US2005022311A1
; GENERAL INFORMATION:
; APPLICANT: APPLERA CORPORATION
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO0315PCT
; CURRENT APPLICATION NUMBER: US/10/504,689
; CURRENT FILING DATE: 2004-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 389627
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(389627)
; OTHER INFORMATION: n = A,T,C or G
US-10-504-689-3

```

	Query Match	4.9%	Score 36.2;	DB 9;	Length 389677;
	Best Local Similarity	59.0%;	Pred. No. 28;		
	Matches	62;	Mismatches 43;	Indels 0;	Gaps 0
QY	102 GGGGCAATTAATGTCTTCGCCAACAACGCTGTATATGATTCCAGAAAACACC CGCGGA	161			

Db 156226 GAGGAAACAGATTCTCTTACAAATGCCATGTTTGTTGTTCTTGAAAAGCACAGCAAAA 156285

Qy 162 CGCATCAACGGTAAAGCGCTACAATTCTTTGGTTAATTGCAGAA 206

Db 156286 CCTTTCAAGCAAACTGTTTCATCATTTCTGGGATTAACACCAAGAA 156330

```

RESULT 3
US-10-322-696-82
; Sequence 82, Application US/10322696
; Publication No. US20040166490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 529452001200
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 405660
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(405660)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-696-82

```

Query Match	4.9%	Score 36.2;	DB 7;	Length 405660;
Best Local Similarity	59.0%	Pred. No. 29;		
Matches 62;	Conservative 0;	Mismatches 43;	Indels 0;	Gaps 0

QY 102 GGGGGCATTAAATGGCTTCGGGAACAACGGTAATGATTCGAGGAAACAACCGGGGAA 161
Db 164226 GAGGGAACGAGATTCCTCTTACAAAGCCATGTTGTGGTCTTGGAAAAGCAGCAAGAA 164285S
QY 163 GGCATCAACGGTAAGCGCTACAACCTCTTGGTAAATTCAGCAAA 206
Db 164286 CCTTCAAGGCAACGTGTTCACTATTGGGATATACACCAAGAA 164330

```

RESULT 4
US-10-487-901-4181/c
; Sequence 4181, Application US/10487901
; Publication No. US20050091708A1
;
GENERAL INFORMATION:
;
APPLICANT: Oreido, Jeremiah Vincent
APPLICANT: McCreery, David
APPLICANT: Pell, Randy
APPLICANT: Miller, Barbara
APPLICANT: Meglartz, Thaddeus
APPLICANT: Gachotte, Daniel
APPLICANT: Blakeslee, Beth
APPLICANT: Larrinua, Ignacio
APPLICANT: Reddy, Avutu
APPLICANT: Shukla, Vipula
APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOM-08552
CURRENT APPLICATION NUMBER: US/10/487,901
CURRENT FILING DATE: 2004-02-26
NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4181

```

```

; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-4181

```

```

Query Match Similarity 4.7%; Score 35.2; DB 9; Length 580;
Best Local Similarity 53.7%; Pred. No. 1.6;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 448 TGGTTATCAATCAATCAATTAAGAAATGACGATCCATCCTATGAGAGTGGTATGAGC 507
    |||||
Db 194 TGGATGATAGTAGTTCCCTTAAACGAGTGAACCACTATTGTAGCCCTTGATGTC 135

Qy 508 ATATTTCATTGTGCAAACTTAAACGGAACTTTATGACCGGATTTACTGCTGAGATA 567
    |||||
Db 134 TTCCTTCCTTTCCCGCACTCATATGATGACCGTCTCTGCAAGCTGCGAAGAGAGAT 75

Qy 568 TATGCAATTAACAATT 583
    |||||
Db 74 TGAAGCAACACAATTT 59

RESULT 5
US-10-369-493-37301
; Sequence 37301, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 37301
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: Thermoplasma volcanium
US-10-369-493-37301

Query Match Similarity 4.7%; Score 35; DB 6; Length 2349;
Best Local Similarity 55.3%; Pred. No. 4.1;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 415 GACGGGATGATGTAACCGGAATGTCGCTTTGGTTATCAACATCAACATAAGAAAT 474
    |||||
Db 1330 GACGGGACAAATGTAACCGAGTGGCCAAATCCGCTCAACCAATTAACAAAAAGGC 1389

Qy 475 GACGATCCACTCTATGAGAGTCTTGAATGAGCATATTCATTGTCGAAACTTAAACGGG 534
    |||||
Db 1390 AAGGTATCTATGCTATATGAGAGATGAAGTGTACATAGACGAATGCTCAACTTAAGCA 1449

Qy 535 AAC 537
    |||
Db 1450 AAC 1452

RESULT 6
US-10-437-963-69024
; Sequence 69024, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

```

; FILE REFERENCE: 38-21 (53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 69024
; LENGTH: 4818
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_69732C.1
US-10-437-963-69024

Query Match
Best Local Similarity 4.6%; Score 34.4; DB 7; Length 4818;
Matches 86; Conservativeness 0; Mismatches 86; Indels 0; Gaps 0;

Oy 558 TGCTGAGTATATATGCAAAATATCAATTCAGGAGCATATTCGACCGACCTCAACCTGGTGT 617
Db 939 TACGAACTAGCTAGGAAAGGAAAGGCGCGCAATATTTAACTGACCAAGTCTTGATGT 998

Oy 618 GTCCATGACGCAACAATCTCGATTTGATATGCGCGATTTTATGACATATGGTCACTGCG 677
Db 999 TTATATGAGGCATATCACTGAGGTGCTGAGAGACAAATATCATCTATATTCGCTCT 1058

Oy 678 GCAGCATTTACTACTTGTGCTGCTTCATGAGGGTTCTTCGCGGTCAAGTAAACG 729
Db 1059 CAAGATTTCTAGGACTTGAACATTTGCGCTGACACCATTTGTAATGAAATG 1110

RESULT 7
US-09-925-065A-724617/c
; Sequence 724617, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 724617
; LENGTH: 1502
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-724617

Query Match
Best Local Similarity 4.6%; Score 34.2; DB 4; Length 1502;
Matches 120; Conservativeness 0; Mismatches 143; Indels 0; Gaps 0;

Oy 320 CGAAGTAGATATATATCCGCTTTCACGACATACGAGGAAGTAGAGAGCTTGAATAATAGAG 379
Db 783 CTATATGAGAGCATTTCTTGCGCATGCTCAAGACCGCAAGAAAGAGCTGTGTGATAGAG 724

Oy 380 TTAATCTCCGCGACGAAGTTAAGATGTGTGGGATACGACGCGATAGGTGAACCGCAATG 439
Db 723 TGAAGTAGATGGGTGTGGAAGGAAGTAGATATATATGTGTGAAGATGTAAACGAGAGTC 664

Oy 440 GTGCGTTTGGGTTATTAACATCACTTAAGAAATAGACGATCACTCTTATGAGAGTGTG 499
Db 663 AGATTAATTAAGGGCTTGAAGAGCCACCATTAAGAAATTTGGCTTTTCTTCAAGTTATATAG 604

```

QY 500 TTGATGGATATTCATTTGAGCACTTAAACGGAGAGCTTATCAGCCGATTTCTACTG 559
|||
Db 603 AAGAGAAATATTCAAATTTTAACTTTTCGTGATGAATTAATAAGAAATCAACA 544
QY 560 CTGAGATATATGCAAAATACAAAT 582
|||
Db 543 AAATGAAATGTACATCTCAATAT 521

RESULT 8

US-09-925-065A-867215
; Sequence 867215, Application US/09925065A
; Publication No. US2005028172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 867215
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-867215

Query Match 4.5%; Score 33.8; DB 4; Length 591;
Best Local Similarity 52.5%; Pred. No. 4.8;

Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 186 TTCTTGGTAAATGCAAGATAGTTCACAGGTAGCAAGCAAGCAAACTCGAGTAC 245
|||
Db 289 TTTTITTTGAGTGGTCTATTTTAAATGTAAATAGCAATACATTAATTTACTAT 348
QY 246 GTCTCCGCTCATAGTGTACCACTCATTCGCAATGCGCAACATCAAGCTGTACGAC 305
|||
Db 349 TTTAAATTTTAAAGTGTACGATTAACAATTTATCATCTTGTGACGCTGTAC 408
QY 306 TTCACTTCTCAGTGTGAAGTT 326
|||
Db 409 TACCATCTCTCTGAACCTT 429

RESULT 9

US-10-425-115-125681
; Sequence 125681, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 125681
; LENGTH: 660

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_46091C.1
US-10-425-115-125681

Query Match 4.5%; Score 33.8; DB 8; Length 660;
Best Local Similarity 52.5%; Pred. No. 5.1;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 584 CAGGACTATTTCCAGACCGACTCAACCTGTGTGTCCATGACGACAAATCTCGATTTG 643
|||
Db 391 CTGGGCACTACTGCACCGAGTACAAAGCCAGGAGTCCAAACGATTTAGAAATGGG 450
QY 644 ATATGCCGAGATTTTATGACATATGTGACGTCGGGCGACATTAATCACTTGTGCTTCA 703
|||
Db 451 TTCAACGGATCCTCTGTGATTTACGTGTGAGAGAAATTAAGTATGATGTCTCTCTA 510
QY 704 TGGGTTTCTTCGGGTTCAGATG 724
|||
Db 511 TGAATGTAGCGCCGACGTACG 531

RESULT 10

US-10-425-114-6745
; Sequence 6745, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6745
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700577561_F11
US-10-425-114-6745

Query Match 4.5%; Score 33.8; DB 7; Length 729;
Best Local Similarity 52.5%; Pred. No. 5.4;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 584 CAGGACTATTTCCAGACCGACTCAACCTGTGTGTCCATGACGACAAATCTCGATTTG 643
|||
Db 464 CTGGGCACTACTGCACCGAGTACAAAGCCAGGAGTCCAAACGATTTAGAAATGGG 523
QY 644 ATATGCCGAGATTTTATGACATATGTGACGTCGGGCGACATTAATCACTTGTGCTTCA 703
|||
Db 524 TTTCACGGATCCTCTGTGATTTACGTGTGAGAGAAATTAAGTATGATGTCTCTCTA 583
QY 704 TGGGTTTCTTCGGGTTCAGATG 724
|||
Db 584 TGAATGTAGCGCCGACGTACG 604

RESULT 11

US-10-425-114-6834
; Sequence 6834, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven B
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO: 6834
LENGTH: 860
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: 700579892_FLI
US-10-425-114-6834

Query Match 4.5%; Score 33.8; DB 7; Length 860;
Best Local Similarity 52.5%; Pred. No. 6;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 584 CAGGACTATTCCGACCGACCTCAACCTGTGTGTCATGACGACCAATCTCGATTG 643
DB 600 CTGGGCATACGTGACCGGAGTACAGCCACGAGATGCGCAACGATGTTAGAAATGG 659
QY 644 ATATGCCGATTTTATGACATATGATGTCACGTGCGGACGACATTCTCACTTGTGCTTCCA 703
DB 660 TTTCAACGGATCTCTGTGATATACGTCTGTGAGAAATTAAGTACGTGCTCTCTCA 719
QY 704 TGGGTTCTTGGGATCAGATG 724
DB 720 TGAAGTAGCCCGACGTACG 740

RESULT 12
US-10-425-115-154424
Sequence 154424, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO: 154424
LENGTH: 935
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MHT4577_72415C.1
US-10-425-115-154424

Query Match 4.5%; Score 33.8; DB 8; Length 935;
Best Local Similarity 52.5%; Pred. No. 6.2;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 584 CAGGACTATTCCGACCGACCTCAACCTGTGTGTCATGACGACCAATCTCGATTG 643
DB 593 CTGGGCATACGTGACCGGAGTACAGCCACGAGATGCGCAACGATGTTAGAAATGG 652
QY 644 ATATGCCGATTTTATGACATATGATGTCACGTGCGGACGACATTCTCACTTGTGCTTCCA 703
DB 653 TTTCAACGGATCTCTGTGATATACGTCTGTGAGAAATTAAGTACGTGCTCTCTCA 712
QY 704 TGGGTTCTTGGGATCAGATG 724
DB 713 TGAAGTAGCCCGACGTACG 733

RESULT 13
US-09-925-065A-41609
Sequence 41609, Application US/09925065A
Publication No. US20050228172A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: PatsSeq for Windows Version 4.0
SEQ ID NO: 41609
LENGTH: 1240
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-41609

Query Match 4.5%; Score 33.8; DB 4; Length 1240;
Best Local Similarity 51.0%; Pred. No. 7.3;
Matches 80; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 484 ACTCCTATGAGAGTGTGATGATGATATTCATTTGAGACTTAAAGGAGCGTTAT 543
DB 145 ACTCTTACTCAGATGCTACTGTTCTGATGACCAAGATTAATACTGAATTTAA 204
QY 544 CAGCCGATTTCTACTGTCGATATATGCAATACAAATTCAGGACATATTCGACGAC 603
DB 205 AATCTCTTCAACCTGTGTGATATATATTAATAAATAATCTGACCTTATACCAA 264
QY 604 CTCACCTGTGTGTCATATGACGACCAATCTCGAT 640
DB 265 GTCAACATTCAGTGAAGATGTCATCAGCAAT 301

RESULT 14
US-10-487-901-3702/c
Sequence 3702, Application US/10487901
Publication No. US20050091708A1
GENERAL INFORMATION:
APPLICANT: Oreido, Jeremiah Vincent
APPLICANT: McCrery, David
APPLICANT: Bell, Randy
APPLICANT: Miller, Barbara
APPLICANT: Weglarz, Thadeus
APPLICANT: Gachotte, Daniel
APPLICANT: Blakeslee, Beth
APPLICANT: Larrinna, Ignacio
APPLICANT: Reddy, Avutu
APPLICANT: Shukla, Vipula
APPLICANT: Crosley, Rodney
TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
FILE REFERENCE: DOW-08552
CURRENT APPLICATION NUMBER: US/10/487,901
CURRENT FILING DATE: 2004-02-26
NUMBER OF SEQ ID NOS: 7560
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 3702
LENGTH: 706
TYPE: DNA
ORGANISM: Artificial

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications_databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_NA_Main) and **.rapbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 22:17:35 ; Search time 531 Seconds
(without alignments)
5605.299 Million cell updates/sec

Title: US-10-784-592-18
Perfect score: 744
Sequence: 1 gtgcgaattatgaagtttt.....aaacgacgtatgctcttcgcg 744

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 200278028 residues
Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /SIDSS/ptodata/1/pubpna/US08 NEW PUB.seq:*
2: /SIDSS/ptodata/1/pubpna/US06 NEW PUB.seq:*
3: /SIDSS/ptodata/1/pubpna/US07 NEW PUB.seq:*
4: /SIDSS/ptodata/1/pubpna/PCT NEW PUB.seq:*
5: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq:*
6: /SIDSS/ptodata/1/pubpna/US09 NEW PUB.seq:*
7: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq:*
8: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq:*
9: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq:*
10: /SIDSS/ptodata/1/pubpna/US10 NEW PUB.seq:*
11: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq:*
12: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq:*
13: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq:*
14: /SIDSS/ptodata/1/pubpna/US11 NEW PUB.seq:*
15: /SIDSS/ptodata/1/pubpna/US60 NEW PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	34.2	4.6	1502	6	US-09-925-065A-724617
c 2	33.8	4.5	591	6	US-09-925-065A-867215
c 3	33.8	4.5	1240	6	US-09-925-065A-41609
c 4	33.8	4.5	1240	9	US-10-301-480-142847
c 5	33.8	4.5	1240	10	US-10-301-480-142847
c 6	33.4	4.5	1848	8	US-10-467-657-1219
c 7	33.2	4.5	534	6	US-09-925-065A-427028
c 8	33.2	4.5	534	10	US-10-301-480-490436
c 9	33.2	4.5	534	10	US-10-301-480-1103845
c 10	33.2	4.5	640	6	US-09-925-065A-400132
c 11	33.2	4.5	643	10	US-10-301-480-467470
c 12	33.2	4.5	643	10	US-10-301-480-1080879
c 13	32.8	4.4	641	6	US-09-925-065A-178736
c 14	32.8	4.4	650	10	US-10-301-480-269706
c 15	32.8	4.4	650	10	US-10-301-480-883115
c 16	32.6	4.4	566	6	US-09-925-065A-604703
c 17	32.6	4.4	566	6	US-09-925-065A-604704
c 18	32.4	4.4	611	6	US-09-925-065A-760902

c 19	32.2	4.3	563	6	US-09-925-065A-192056
c 20	32.2	4.3	571	10	US-10-301-480-281632
c 21	32.2	4.3	571	10	US-10-301-480-895041
c 22	32.2	4.3	598	6	US-09-925-065A-301137
c 23	32.2	4.3	633	6	US-09-925-065A-914168
c 24	31.8	4.3	528	10	US-10-301-480-430289
c 25	31.8	4.3	528	10	US-10-301-480-1043698
c 26	31.8	4.3	533	6	US-09-925-065A-413580
c 27	31.8	4.3	551	10	US-10-301-480-479582
c 28	31.8	4.3	551	10	US-10-301-480-1092991
c 29	31.8	4.3	557	10	US-10-301-480-492045
c 30	31.8	4.3	557	10	US-10-301-480-1105454
c 31	31.8	4.3	558	6	US-09-925-065A-359103
c 32	31.8	4.3	573	6	US-09-925-065A-429103
c 33	31.8	4.3	615	6	US-09-925-065A-517010
c 34	31.8	4.3	1400	14	US-11-136-527-6130
c 35	31.8	4.3	4953	14	US-11-136-527-2034
c 36	31.6	4.2	633	14	US-11-000-463-104
c 37	31.6	4.2	634	14	US-11-000-463-576
c 38	31.6	4.2	1394	6	US-09-925-065A-711186
c 39	31.4	4.2	540	6	US-09-925-065A-747461
c 40	31.4	4.2	545	6	US-09-925-065A-746148
c 41	31.4	4.2	611	6	US-09-925-065A-226317
c 42	31.2	4.2	505	10	US-10-301-480-471976
c 43	31.2	4.2	505	10	US-10-301-480-1085385
c 44	31.2	4.2	524	6	US-09-925-065A-405101
c 45	31.2	4.2	1173	8	US-10-517-939-27

ALIGNMENTS

RESULT 1

US-09-925-065A-724617/c
; Sequence 724617, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2001-08-08
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 724617
; LENGTH: 1502
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-724617

Query Match 4.6%; Score 34.2; DB 6; Length 1502;
Best Local Similarity 45.6%; Pred. No. 3.5;
Matches 120; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy	320	CGAAGTTACGATATATCCGTTTTCACATACGGAGGTTAGGACTTGGAAATTAGAG	379
Dy	783	CTAATGTAGGAGCATTTCTGCGATGCTCAAAGACCAGCAAGAGCTGTGTGATAGAG	724
Qy	380	TTAATCTCTGCGAGCAAGTTAAGAGTGTGGGTAGCGGGTAGGTGAACCCGAATG	439
Dy	723	TGAAGTGGATGGGTGTGTGAAAGGAAGTAGTATATATGTGGTAAAGATGTATACAGGAGGTC	664

Qy	440	GTGCGTTTTGGTTATCAACATCACCATAGAAATGACGGATCCACTCTTATCGAGGTCTG	499
Db	663	AGATAATTAGGCTTTTGAAGGCCACCAATAGAAATTTGGCTTTTCTTCAAGTTTATATAG	604
Qy	500	TTGATGGCATATTCOATTTGTCAGAACTTAAACGGGAACGTTTATCAGCCGGAATCTTACTG	559
Db	603	AGAAGAATATTCAAATTTTTTAAACTTTTCTGTCATGAATTATTAATAAAGAATACATGA	544
Qy	560	CTGAGATATATCGAAATACAAAT	582
Db	543	AAATGAATGTACATCTCAATAT	521

RESULT 2

```

US-09-925-065A-867215
; Sequence 867215, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 867215
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-867215

```

RESULT 3

```

US-09-925-065A-41609
; Sequence 41609, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147

```

```

; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41609
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-41609

```

RESULT 4

```

US-10-301-480-142847
; Sequence 142847, Application US/10301480
; Publication No. US20060057564A1
;
; GENERAL INFORMATION:
;
;   APPLICANT: Wang, David G.
;
;   TITLE OF INVENTION: Identification and Mapping of
;   TITLE OF INVENTION: in the Human Genome
;
;   FILE REFERENCE: 108827.137
;
;   CURRENT APPLICATION NUMBER: US/10/301,480
;   CURRENT FILING DATE: 2002-11-21
;   PRIOR APPLICATION NUMBER: US 10/215,598
;   PRIOR FILING DATE: 2002-08-09
;   PRIOR APPLICATION NUMBER: US 60/311,695
;   PRIOR FILING DATE: 2001-08-10
;   NUMBER OF SEQ ID NOS: 1226818
;
; SOFTWARE: fastSEQ for Windows Version 4.0
;
; SEQ ID NO 142847
;
;   LENGTH: 1240
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-142847

```

RESIST. 5

```
US-10-301-480-756256
; Sequence 756256, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 756256
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-756256

Query Match          4.5%; Score 33.8; DB 10; Length 1240;
Best Local Similarity 51.0%; Pred. No. 4.4;
Matches 80; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 484 ACTCGATGAGGTCGTTGATGGCATATTCATTTGCAGAACTTAACCGGGAACGTTTAT 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 ACTCTTTTACCTCAGATGCTACTGTTCTGATGACCAAGATTAATAACTGAACCTTAAA 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 544 CAGCGGGAATTCATGCTGAGATATATGCAATACAAATTCAGGACTATTCCGACCGAC 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 AATCTCTCTTCACCTGTGTATATATATAAAAAAAATACTGACCTTTTATCACCAAA 264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 604 CTCACCTCGTGTGTCATGACGACAAATCTCGTAT 640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 GTCAACAATCCAGTGAAGAAAGATGGTCACATCAGGAAT 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-10-467-657-1219
; Sequence 1219, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1219
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1219

Query Match          4.5%; Score 33.4; DB 8; Length 1848;
Best Local Similarity 65.3%; Pred. No. 7.1;
Matches 49; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 122 CGAACACGGTGATATGATTCGAGAAACACACCGCGGACGCATCAACGGTAAGCGCTA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1281 CGACACGGTTTTTACGATATGGCGGTCAACAGCGCGGACCGACCGCAAACTCAA 1340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 182 CAACTTCTTTGGTTA 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 1341 CTACACCTTGCGCTA 1355

RESULT 7
US-09-925-065A-427028/c
; Sequence 427028, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925.065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427028
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-427028

Query Match          4.5%; Score 33.2; DB 6; Length 534;
Best Local Similarity 52.2%; Pred. No. 4.6;
Matches 71; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

Qy 376 AGAGTTAACTCCCTGCAGCAAGTTAAGAGTGTGGGGTACGACGGGATAGGTGAACCGCA 435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 ACATTTAACAGGCTGCAAGAAGTTGCTCTCTCTCCCACTGGGGTAGGTCAGGACT 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 436 AATGGTGGTTTTGGGTTATCAATCACCATAAGAATGACGGATCCATCTCTATGGAG 495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 CACAGAGAGTCAGGCGCTTTTCATCATTTACACAGAGTAAGAGCCATTCACCCCATGTG 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 496 GTCGTTGATGCGCATAT 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 CCTGTGTGGCCATGT 56
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-301-480-490436/c
; Sequence 490436, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 490436
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-490436

Query Match          4.5%; Score 33.2; DB 10; Length 534;
```

```
Best Local Similarity 52.2%; Pred. No. 4.6;
Matches 71; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

Qy 376 AGAGTTAACTCCCTGCAGCAAGTTAAGAGTGTGGGTACGACGGGATAGGTGAACCGCA 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 ACATTTAACAGGCTGCAAGAAGTTCCTGCTCTCCCAATGGGGTAGGTGAGGACT 132
Qy 436 AATGTCGGTTCCTGGTTATCAATCACCATAGAAATGACGGATCCCTCTATGGAG 495
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 CACAGAGTGCAGGCTTTTCATCATTAACAGAGTAAAGAGCCATTACCCCCCATGGT 72
Qy 496 GTCGTTCATGCGCATAT 511
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 CCTGTGTGGCCATGT 56

RESULT 9
US-10-301-480-1103845/c
; Sequence 1103845, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1103845
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1103845

Query Match 4.5%; Score 33.2; DB 10; Length 534;
Best Local Similarity 52.2%; Pred. No. 4.6;
Matches 71; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

Qy 376 AGAGTTAACTCCCTGCAGCAAGTTAAGAGTGTGGGTACGACGGGATAGGTGAACCGCA 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 ACATTTAACAGGCTGCAAGAAGTTCCTGCTCTCCCAATGGGGTAGGTGAGGACT 132
Qy 436 AATGTCGGTTCCTGGTTATCAATCACCATAGAAATGACGGATCCCTCTATGGAG 495
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 CACAGAGTGCAGGCTTTTCATCATTAACAGAGTAAAGAGCCATTACCCCCCATGGT 72
Qy 496 GTCGTTCATGCGCATAT 511
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 CCTGTGTGGCCATGT 56

RESULT 10
US-09-925-065A-400132/c
; Sequence 400132, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
```

```
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400132
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-400132

Query Match 4.5%; Score 33.2; DB 6; Length 640;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 50; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

Qy 2 TGGCAATTATGAAAGTTTGGGATGGATTTTGGTACCGTATATCATGCTGTTTATTCAGT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
313 TTCTTATATTCAAAGTATTGACTTTGTATCTTTGTACCTTATATTGATGTTTATTCAGT 254
Qy 62 GGGGGCGAATGAACAGAAATT 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GTGTTTCAAATCAAAGTTT 234

RESULT 11
US-10-301-480-467470/c
; Sequence 467470, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467470
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-467470

Query Match 4.5%; Score 33.2; DB 10; Length 643;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 50; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

Qy 2 TGGCAATTATGAAAGTTTGGGATGGATTTTGGTACCGTATATCATGCTGTTTATTCAGT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
313 TTCTTATATTCAAAGTATTGACTTTGTATCTTTGTACCTTATATTGATGTTTATTCAGT 254
Qy 62 GGGGGCGAATGAACAGAAATT 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GTGTTTCAAATCAAAGTTT 234

RESULT 12
US-10-301-480-1080879/c
; Sequence 1080879, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 60/311,695
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1080879
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1080879

Query Match          4.4%; Score 33.2; DB 10; Length 643;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 50; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

Qy 2 TCGCAATTATGAAAGTTTGGGATCGATTTTGGTACCGTATATCATCTGTTTATTCAGT 61
Db 313 TTCTATATTCAAGTATGACTTGTACTTTGTACCTTATATTGATGTTTATTCACT 254

Qy 62 GGGGCGGAATGAACGAATT 81
Db 253 GTGTTTCAATCAAGTTT 234

RESULT 13
US-09-925-065A-178736/c
; Sequence 178736, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178736.1
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-178736

Query Match          4.4%; Score 32.8; DB 6; Length 641;
Best Local Similarity 55.2%; Pred. No. 6.8;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 299 CTAGGACTTCATCTTCTCAGTCGAGTTACGATATATCCCGTTTTCACACATACGGGAAGG 358
Db 618 CTTTGCTTCATTTTTTAAATATAAATACTAGGATATCACTTGTTCCTTTGCTTCTTGG 559

Qy 359 TAGGAGACTTGGAAATTAGAGTTAACTCCCTGCAGCAAGTTTAAGAGTGTGGGTAC 414
Db 558 TATAGCGCTGTAAACTTACATAAATTTCCAGCATTTCAGTTTAAGCTGTGCGATAC 503

RESULT 14
US-10-301-480-269706/c
; Sequence 269706, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
```

```
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269706
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-269706

Query Match          4.4%; Score 32.8; DB 10; Length 650;
Best Local Similarity 55.2%; Pred. No. 6.8;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 299 CTAGGACTTCATCTTCTCAGTCGAGTTACGATATATCCCGTTTTCACACATACGGGAAGG 358
Db 618 CTTTGCTTCATTTTTTAAATATAAATACTAGGATATCACTTGTTCCTTTGCTTCTTGG 559

Qy 359 TAGGAGACTTGGAAATTAGAGTTAACTCCCTGCAGCAAGTTTAAGAGTGTGGGTAC 414
Db 558 TATAGCGCTGTAAACTTACATAAATTTCCAGCATTTCAGTTTAAGCTGTGCGATAC 503

RESULT 15
US-10-301-480-883115/c
; Sequence 883115, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 883115
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-883115

Query Match          4.4%; Score 32.8; DB 10; Length 650;
Best Local Similarity 55.2%; Pred. No. 6.8;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 299 CTAGGACTTCATCTTCTCAGTCGAGTTACGATATATCCCGTTTTCACACATACGGGAAGG 358
Db 618 CTTTGCTTCATTTTTTAAATATAAATACTAGGATATCACTTGTTCCTTTGCTTCTTGG 559

Qy 359 TAGGAGACTTGGAAATTAGAGTTAACTCCCTGCAGCAAGTTTAAGAGTGTGGGTAC 414
Db 558 TATAGCGCTGTAAACTTACATAAATTTCCAGCATTTCAGTTTAAGCTGTGCGATAC 503

Search completed: April 7, 2006, 22:41:16
Job time : 533 secs
```

This page blank (uspro)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: April 7, 2006, 16:08:38 ; Search time 295.965 Seconds
(without alignments)
368.172 Million cell updates/sec

Title: US-10-784-592-43

Perfect score: 1289
Sequence: 1 KRIMKVLGILIVYIMLFIQ.....YSLVSMGPFSGSDFTTYALP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_21: *
1: geneeqp19808: *
2: geneeqp19908: *
3: geneeqp20008: *
4: geneeqp20018: *
5: geneeqp20028: *
6: geneeqp20038: *
7: geneeqp20048: *
8: geneeqp20058: *
9: geneeqp20068: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1289	100.0	248	9	AEb45569 Alicyclob
2	1289	100.0	248	9	AEb48755 Alicyclob
3	1074	83.3	207	9	AEb45583 Alicyclob
4	135	10.5	331	4	AAU37553 Staphyloc
5	135	10.5	331	4	AAU34416 Staphyloc
6	135	10.5	331	4	AAU37279 Staphyloc
7	135	10.5	331	4	ABU16122 Protein e
8	135	10.5	331	6	ABU15887 Protein e
9	113.5	8.8	6713	6	ABU15887 Protein e
10	107.5	8.3	2086	4	AAU34143 Staphyloc
11	107.5	8.3	5795	4	AAU37017 Staphyloc
12	106.5	8.3	417	7	ABO70979 Pseudomon
13	106	8.2	2344	4	AAU37120 Staphyloc
14	105.5	8.0	337	5	ADJ34822 Xylanae
15	105.5	8.0	337	5	AAE19444 Hansenula
16	102	7.9	2261	6	ABU18914 Pathogen
17	102	7.9	2271	6	ABU16000 Protein e
18	102	7.9	2271	6	ABU16000 Protein e
19	102	7.9	2283	6	ABU56876 Staphyloc
20	100	7.8	982	4	ABU59353 Drosophila
21	100	7.8	982	4	ABU67401 Drosophila
22	99	7.7	194	6	ABU29785 Protein e
23	99	7.7	485	6	ABP56884 Staphyloc
24	96.5	7.6	2398	6	ABU42252 Protein e

25	98.5	7.6	6281	4	AAU37403	AAU37403 Staphyloc
26	98.5	7.6	9535	6	ABM73008	ABM73008 Staphyloc
27	98.5	7.6	10498	6	ABJ19119	ABJ19119 Pathogen
28	96.5	7.5	910	3	AAE25104	AAE25104 Pinus rad
29	96	7.4	2233	8	ADG43906	ADG43906 Human DP4
30	95.5	7.4	632	7	ADP60974	ADP60974 B. thurin
31	95.5	7.4	786	8	ADS27302	ADS27302 Bacterial
32	95.5	7.4	791	8	ADS26563	ADS26563 Bacterial
33	95.5	7.4	805	8	ADS26940	ADS26940 Bacterial
34	94.5	7.3	458	6	ABU26050	ABU26050 Protein e
35	94.5	7.3	498	1	AAU70302	AAU70302 Sequence
36	94.5	7.3	522	7	ADG65590	ADG65590 B. faeciu
37	94.5	7.3	627	7	ADP60976	ADP60976 B. thurin
38	94	7.3	349	3	AAE26877	AAE26877 Glucosylt
39	94	7.3	829	8	ADU99233	ADU99233 Environme
40	94	7.3	1375	5	AAU98028	AAU98028 S. mutans
41	94	7.3	1375	5	AAU9288	AAU9288 Streptoco
42	94	7.3	1375	5	ADU93655	ADU93655 Streptoco
43	94	7.3	1375	9	ADU37278	ADU37278 Streptoco
44	94	7.3	1484	8	ADJ34858	ADJ34858 Xylanae
45	94	7.3	2712	6	ABU39146	ABU39146 Protein e

ALIGNMENTS

RESULT 1	AEb45569	standard; protein; 248 AA.
XX	AEb45569;	
AC	AEb45569;	
DT	22-SEP-2005	(first entry)
XX		
DR	Alicyclobacillus sp. functional polypeptide, SEQ ID NO: 43.	
XX		
KW	Feedstuff; food; detergent; surfactant; pulp; functional polypeptide.	
XX		
OS	Alicyclobacillus sp.; DSM 15716.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..41
FT	Misc-difference	/label= Signal_peptide
FT	Protein	/note= "Encoded by GTC"
FT		42..248
FT		/note= "Alicyclobacillus sp. mature functional polypeptide"
XX		
PN	US2005147983-A1.	
XX		
PD	07-JUL-2005.	
XX		
PF	23-FEB-2004; 2004US-00784592.	
XX		
PR	06-JAN-2004; 2004DK-00000010.	
XX		
PR	04-FEB-2004; 2004DK-00000165.	
XX		
PA	(NOVO) NOVOZYMES AS.	
XX		
PI	Wilting R, Laasen SF, Ostergaard PR,	
XX		
DR	WPI; 2005-511773/52.	
XX		
DR	N-PSDB; AEB45544.	
XX		
PT	New functional polypeptides having function and amino acid sequence	
XX		
PT	similar to known specific bacterial enzymes useful in industrial,	
XX		
PT	research and household applications e.g. detergents and food.	
XX		
PS	Example 2; SEQ ID NO 43; 83bp; English.	
XX		
CC	The present invention relates to functional polypeptides encoded by	
XX		
CC	polynucleotides comprised in the genome of Alicyclobacillus sp. DSM	

CC 15716. The invention is useful in industrial, research and household
CC processes such as in composition to prepare feed and food, in detergent
CC formulations and for treating lignocellulosic fabric and pulp. The
CC present sequence is Alicyclobacillus sp. DSM 15716 functional
CC polypeptide.

XX Sequence 248 AA;

Query Match 100.0%; Score 1289; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 3e-108;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVPANTYMRGNTPRNASTVSA 60
DB 1 MRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVPANTYMRGNTPRNASTVSA 60
QY 61 TTSLVNSTNSSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSGLRYIPFTYGV 120
DB 61 TTSLVNSTNSSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSGLRYIPFTYGV 120
QY 121 GDLEIRVNSLOOVKSVGVDIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180
DB 121 GDLEIRVNSLOOVKSVGVDIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180
QY 181 YQDSTAEIYANTNSGIIPTDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240
DB 181 YQDSTAEIYANTNSGIIPTDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240
QY 241 DDTYALP 248
DB 241 DDTYALP 248

RESULT 2

ID AEB48755
AEB48755 standard; protein, 248 AA.

AC AEB48755;

DT 22-SEP-2005 (first entry)

DE Alicyclobacillus sp. DSM 15716 functional polypeptide.

KM Protein production; protein secretion; surfactant; feedstuff; food.

OS Alicyclobacillus sp.; 'DSM 15716'.

FT Key Location/Qualifiers

FT Peptide 1..41

FT Misc-difference 1/label= Signal_peptide

FT Protein 42..248/note= "Encoded by GTC"

FT 42..248

FT 42..248

FT 42..248

FT 42..248

FT 42..248

FT 42..248

FT 42..248

FT 42..248

FT 42..248

FT 42..248

PT New mature functional polypeptide of Alicyclobacillus sp., used for
PT commercial research purposes, specifically for cleaning a cellulosic
PT fabric, preparing food or feed additive, and for treating lignocellu-
PT materials and pulp.

PS Claim 3; SEQ ID NO 43; 151pp; English.

XX The present invention is based on the finding of a strain of
CC Alicyclobacillus, namely Alicyclobacillus sp. DSM 15716, which grows at
CC low pH (approximately 4-5) and at high temperature (50-60 degrees C). It
CC is an object of the invention to identify and provide polypeptides
CC secreted from Alicyclobacillus sp. DSM 15716 because such polypeptides
CC may be used for industrial purposes and may also be produced in
CC industrially relevant processes and amounts. Thus, the invention provides
CC isolated mature functional polypeptides which are at least 90% identical
CC to, and exhibit and same function as a corresponding secreted polypeptide
CC AEB48738-AEB48762 from Alicyclobacillus sp. DSM 15716, and also provides
CC polynucleotides AEB48713-AEB48737 encoding such polypeptides. The
CC polypeptides are preferably enzymes having acid endoglycanase, acid
CC cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl
CC isomerase, gamma-D-glutamyl-L-lysine aminopeptidase, endo-beta-N-
CC acetylglucosaminidase, peptidyl-prolyl-isomerase, xylanase
CC phytase, phospholipase C, polysaccharide deacetylase, xylanase
CC deacetylase or sulfite oxidase activity. A composition comprising such an
CC enzyme can be used in a detergent composition, or a food or feed
CC composition. Also claimed are: nucleic acid constructs, recombinant
CC expression vectors and host cells; methods of producing the polypeptides;
CC a storage medium suitable for use in an electronic device comprising
CC information of the amino acid sequence of a polypeptide of the invention
CC or the nucleotide sequence encoding the polypeptide; and a process
CC comprising employing the polypeptide, or polynucleotide encoding it, in
CC an industrial or household technical process. The present sequence is the
CC protein sequence of a functional polypeptide of the invention.

XX Sequence 248 AA;

Query Match 100.0%; Score 1289; DB 9; Length 248;
Best Local Similarity 100.0%; Pred. No. 3e-108;
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVPANTYMRGNTPRNASTVSA 60
DB 1 MRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVPANTYMRGNTPRNASTVSA 60
QY 61 TTSLVNSTNSSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSGLRYIPFTYGV 120
DB 61 TTSLVNSTNSSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSGLRYIPFTYGV 120
QY 121 GDLEIRVNSLOOVKSVGVDIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180
DB 121 GDLEIRVNSLOOVKSVGVDIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180
QY 181 YQDSTAEIYANTNSGIIPTDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240
DB 181 YQDSTAEIYANTNSGIIPTDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240
QY 241 DDTYALP 248
DB 241 DDTYALP 248

RESULT 3

ID AEB45583
AEB45583 standard; protein, 207 AA.

AC AEB45583;

DT 22-SEP-2005 (first entry)

DE Alicyclobacillus sp. mature functional polypeptide (amino acids 42-248).

KM Feedstuff; food; detergent; surfactant; pulp; functional polypeptide.

```
XX OS Alicyclobacillus sp.; DSM 15716.
XX PN US2005147983-牌..
XX PD 07-JUL-2005.
XX PF 23-FEB-2004; 2004US-00784592.
XX PR 06-JAN-2004; 2004DK-00000010.
XX PR 04-FEB-2004; 2004DK-00000165.
XX PA (NOVO ) NOVOZYMES AS.
XX PI Wilting R, Lassen SF, Ostergaard PR;
XX DR WPI; 2005-511773/52.
XX DR N-PSDB; AEB45582.
XX PT New functional polypeptides having function and amino acid sequence
XX PT similar to known specific bacterial enzymes useful in industrial,
XX PT research and household applications e.g. detergents and food.
XX PS Claim 36; Page; 83pp; English.
XX CC The present invention relates to functional polypeptides encoded by
XX CC polynucleotides comprised in the genome of Alicyclobacillus sp. DSM
XX CC 15716. The invention is useful in industrial, research and household
XX CC processes such as in composition to prepare feed and food, in detergent
XX CC formulations and for treating lignocellulosic fabric and pulp. The
XX CC present sequence is Alicyclobacillus sp. DSM 15716 mature functional
XX CC polypeptide. Note: this sequence is not shown in the specification but is
XX CC constructed based on the amino acid positions provided in claim 36 of the
XX CC specification.
XX SQ Sequence 207 AA;
Query Match 83.3%; Score 1074; DB 9; Length 207;
Best Local Similarity 100.0%; Pred. No. 7e-89;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 NTVVMKRGTPRNASTVATSLVNSTNSQVAKOEONSTSPAKHSTNSLQHAHQAT 101
DB 1 NTVVMKRGTPRNASTVATSLVNSTNSQVAKOEONSTSPAKHSTNSLQHAHQAT 60
QY 102 TSSSQSKLRYIPFHTYGVGDELEIRVNSIQVKSQVYDGIGETANGAFVINITIRNDG 161
DB 61 TSSSQSKLRYIPFHTYGVGDELEIRVNSIQVKSQVYDGIGETANGAFVINITIRNDG 120
QY 162 TMEVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTTPTDLNPGVSMTNLVFMDPDMT 221
DB 121 TMEVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTTPTDLNPGVSMTNLVFMDPDMT 180
QY 222 YGVHGOHYSLVASMGFSGSDETTYALP 248
DB 181 YGVHGOHYSLVASMGFSGSDETTYALP 207
RESULT 4
AAU37553
ID AAU37553 standard; protein; 331 AA.
XX AC AAU37553;
XX DT 14-FEB-2002 (first entry)
XX DB Staphylococcus aureus cellular proliferation protein #1723.
XX CC Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX KW antibacterial; drug design.
XX OS Staphylococcus aureus.
```

```
PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0263088P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS5412.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 13146; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes,
XX CC themselves and the encoded proteins. The prokaryotes used are Bacterichia
XX CC coili, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 AA;
Query Match 10.5%; Score 135; DB 4; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.0016;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
QY 115 HTYGVGDELEIRVNSIQVKSQVYDGIGETANGAFVINITIRNDGSTMEVVDGIF 171
DB 81 HRTGTNKGDELVVNSVETWKSVPGLAPNAGTFVADVTKKNGKALTTDSSMF 140
QY 172 HLQNLNGVYQPDSTAEIYAN-TNSGTTPTD-----LNPGVSMTNLVFMDPMD 218
DB 141 KXKS-GDTPFADNTGMSANGSDNGSIENSFFLQRIKPNDSRAQKIVFDSB 192
RESULT 5
AAU34416
ID AAU34416 standard; protein; 331 AA.
XX AC AAU34416;
XX DT 14-FEB-2002 (first entry)
XX DB Staphylococcus aureus cellular proliferation protein #692.
XX CC Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX KW antibacterial; drug design.
```

```
XX OS Staphylococcus aureus.
XX PN MO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX P1 Haeselbeck R, Ohlsen KU, Zysek JD, Wall D, Trawick JD, Carr GJ,
XX P1 Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS52275.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 5912; 511pp; English.
XX XX The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes,
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 AA;
XX
XX Query Match 10.5%; Score 135; DB 4; Length 331;
XX Best Local Similarity 31.0%; Pred. No. 0.0016;
XX Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
OY 115 HTYG---KVGDLIRVNSLQGVKSVGDIGETANGAFWVNIITIRNDGTPMEVVDGIF 171
DB 81 HKIGTVNGDLRVTNVSVETKSVGSPSLAPNAGIFVADVITIKKKGKALTTIDSSMF 140
OY 172 HLOINGNVYQPDSTAETIYAN-TNSGTIPTD-----LNPQVSMNTNLVFDMPD 218
DB 141 KTKS-GDKTFEADNTGMSANQSDNGSIENSFFLGRINPDSTAQKIVFVDSR 192
XX
XX RESULT 6
XX ID AAU37279 standard; protein, 331 AA.
XX AC AAU37279;
XX XX 14-FEB-2002 (first entry)
XX XX Staphylococcus aureus cellular proliferation protein #1449.
XX DE
```

```
XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX KW antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN MO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX P1 Haeselbeck R, Ohlsen KU, Zysek JD, Wall D, Trawick JD, Carr GJ,
XX P1 Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS55138.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 12672; 511pp; English.
XX XX The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 AA;
XX
XX Query Match 10.5%; Score 135; DB 4; Length 331;
XX Best Local Similarity 31.0%; Pred. No. 0.0016;
XX Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
OY 115 HTYG---KVGDLIRVNSLQGVKSVGDIGETANGAFWVNIITIRNDGTPMEVVDGIF 171
DB 81 HKIGTVNGDLRVTNVSVETKSVGSPSLAPNAGIFVADVITIKKKGKALTTIDSSMF 140
OY 172 HLOINGNVYQPDSTAETIYAN-TNSGTIPTD-----LNPQVSMNTNLVFDMPD 218
DB 141 KTKS-GDKTFEADNTGMSANQSDNGSIENSFFLGRINPDSTAQKIVFVDSR 192
XX
XX RESULT 7
XX ID ABU16122 standard; protein, 331 AA.
XX AC ABU16122;
XX XX
```


XX AC ABU15887;
 XX 19-JUN-2003 (first entry)
 XX DB Protein encoded by Prokaryotic essential gene #1414.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX OS Staphylococcus aureus.
 XX PN W020027183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002MO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 XX PR 06-SEP-2001; 2001US-00948933.
 XX PR 25-OCT-2001; 2001US-0342923P.
 XX PR 08-FEB-2002; 2002US-00072851.
 XX PR 06-MAR-2002; 2002US-0362693P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW,
 XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX DR WPI; 2003-029926/02.
 XX DR N-PDB; ACN19757.
 XX PT New antisense nucleic acids, useful for identifying proteins or screening
 XX PT for homologous nucleic acids required for cellular proliferation to
 XX PT isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 43811; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 XX CC the 6213 antisense sequences given in the specification where expression
 XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
 XX CC encoding a polypeptide whose expression is inhibited by the antisense
 XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX CC polypeptide or its fragment whose expression is inhibited by the
 XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
 XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX CC proliferation or the activity of a gene in an operon required for
 XX CC proliferation; (7) identifying a compound that influences the activity of
 XX CC the gene product or that has an activity against a biological pathway
 XX CC required for proliferation, or that inhibits cellular proliferation; (8)
 XX CC identifying a gene required for cellular proliferation or the biological
 XX CC pathway in which a proliferation-required gene or its gene product lies
 XX CC or a gene on which the test compound that inhibits proliferation of an
 XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX CC compound's activity; (11) a culture comprising strains in which the gene
 XX CC product is overexpressed or underexpressed; (12) determining the extent
 XX CC to which each of the strains is present in a culture or collection of
 XX CC strings; or (13) identifying the target of a compound that inhibits the
 XX CC proliferation of an organism. The antisense nucleic acids are useful for
 XX CC identifying proteins or screening for homologous nucleic acids required
 XX CC for cellular proliferation to isolate candidate molecules for rational
 XX CC drug discovery programs, or for screening homologous nucleic acids
 XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 XX CC the target prokaryotic essential genes. Note: The sequence data for this
 XX CC patent did not form part of the printed specification, but was obtained
 XX CC in electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 6713 AA;

Query Match 8.8%; Score 113.5; DB 6; Length 6713;
 Best Local Similarity 24.4%; Pred. No. 8.3;

Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;
 QY 42 NTVMIRGNTPRNAST--VSATTSILVNSTNSQVAKQGNSTSPRAHSTNSLQAHQQA 99
 DB 899 NONTYLDADSRNRYTQAVTAABGILNKQTCGNTSKADVNALNAVTAKAALNGASRLR 958
 QY 100 ATSSSOSKLRVIFPHTYTGKGDLEIRVNSLQGVKSVGYDGI---GETANGAFWYINITY 156
 DB 959 NAKTSATNTINGLNLNLOLOKDNLKHQVEQAQNV--YGVNVKDKGNTLNTRMGALRTSI 1016
 QY 157 RNDSTPMNEVVDGIFHLQNLNGVNYQPDSTAEIYANTNSGTIPDLNPGVMTNLVPRDM 216
 DB 1017 QNDWYTT--KTSQNTYLDASDSKNNY---NTA---VNNANGVYINATNND--NMDANAINDM 1066
 RESULT 10
 AAU34143
 ID AU34143 standard; protein; 2086 AA.
 AC AAU34143;
 DT 14-FEB-2002 (first entry)
 DE Staphylococcus aureus cellular proliferation protein #419.
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 KW antibacterial; drug design.
 OS Staphylococcus aureus.
 PN W0200170955-A2.
 PD 27-SEP-2001.
 PF 21-MAR-2001; 2001WO-US009180.
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 PA (ELIT-) ELITRA PHARM INC.
 PI Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ,
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 DR N-PDB; AAS52002.
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 PS Example 3; SEQ ID NO 5639; 511pp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 XX CC prokaryotic cellular proliferation, their use in identifying the genes,
 XX CC their use in the discovery of novel antibiotics, the essential genes
 XX CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 XX CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 XX CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 XX CC useful for the identification of potential new targets for antibiotic
 XX CC development. The antisense nucleic acids can also be used to identify
 XX CC proteins used in proliferation, to express these proteins, and to obtain
 XX CC antibodies capable of binding to the expressed proteins. The proteins can
 XX CC be used to screen compounds in rational drug discovery programmes. The
 XX CC antisense nucleic acid sequence is also useful to screen for homologous
 XX CC nucleic acids which are required for cell proliferation in a wide variety
 XX CC of organisms. The present sequence represents an essential prokaryotic
 XX CC cellular proliferation protein. Note: The sequence data for this patent
 XX CC did not form part of the printed specification, but was obtained in

```
CC electronic format directly from WIP0 at
CC ftp.wipo.int/pub/published_pct_sequences
CC
CC Sequence 2086 AA;
SQ
Query Match      8.3%; Score 107.5; DB 4; Length 2086;
Best Local Similarity 24.2%; Pred. No. 5.9;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;
Qy
Db 42 NTVMIRGNTPRMAST--VSATTSLVNSTNSQVAKQKQNSSTSPAKHSTNSLQHQHQ 99
1323 NQNYLDADESKRNATQAVTAABGILNKQGTGNTSKADVDAALMTVTAKAALNGAEHLR 1382
Qy 100 ATSSSOSKLRYPHTYGVKVDLEIRVNSLQVQKSV-GYDGIETANGAFWVINITRN 158
Db 1383 NTKTSATNTINGLPLNLQLOKDNLKHQVEQKQVAGVGVDKGNTLNTANGALRTSIQ 1442
Qy 159 DGSTPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTPTDLNPGVSMTTNLVFD 216
Db 1443 DNTT--KTSQNYLDASDSNKKNY---NTA---VNNANGVIVNTNPN--NMDANALNGM 1490
RESULT 11
ID AU37017 standard; protein; 5795 AA.
XX
AC AU37017;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1187.
XX
KM Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (BLIT-) ELITRA PHARM INC.
XX
PI Haselebeck R, Ohlsen KU, Zyekind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
XX N-PSDB; AAS54876.
XX
XX
XX New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX
XX Example 3; SEQ ID NO 12610; 511P; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Bacteri-
XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
```

```
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIP0 at
CC ftp.wipo.int/pub/published_pct_sequences
CC
CC Sequence 5795 AA;
SQ
Query Match      8.3%; Score 107.5; DB 4; Length 5795;
Best Local Similarity 24.2%; Pred. No. 24;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;
Qy
Db 42 NTVMIRGNTPRMAST--VSATTSLVNSTNSQVAKQKQNSSTSPAKHSTNSLQHQHQ 99
3600 NQNYLDADESKRNATQAVTAABGILNKQGTGNTSKADVDAALMTVTAKAALNGAEHLR 3659
Qy 100 ATSSSOSKLRYPHTYGVKVDLEIRVNSLQVQKSV-GYDGIETANGAFWVINITRN 158
Db 3660 NTKTSATNTINGLPLNLQLOKDNLKHQVEQKQVAGVGVDKGNTLNTANGALRTSIQ 3719
Qy 159 DGSTPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTPTDLNPGVSMTTNLVFD 216
Db 3720 DNTT--KTSQNYLDASDSNKKNY---NTA---VNNANGVIVNTNPN--NMDANALNGM 3767
RESULT 12
ID ABO70979 standard; protein; 417 AA.
XX
AC ABO70979;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #3154.
XX
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US651795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX
XX N-PSDB; ABD04550.
XX
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX
XX
XX Disclosure; SEQ ID NO 19725; 4559P; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
```


CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology Sequences ABO67826-
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 417 AA;

Query Match 8.3%; Score 106.5; DB 7; Length 417;
Best Local Similarity 25.6%; Pred. No. 0.82;
Matches 58; Conservative 33; Mismatches 79; Indels 57; Gaps 12;

QY 32 GSWALIVFANTVVMIRGNTPRNASTVSATSLVNST-----TSLSVNS----- 67
DB 21 GIKKALIVNTNTIASL-----NTGRNLNNSASALNTSLQRLSTGRINSKADDAAGLIQIANRL 77
QY 68 ---TNSQVAKQEQNSSTSPA-----HKSTNSLQHAQH---QAATTSQSQSLRYIPF 114
DB 78 TSGVNGALVNTKNNANDGISLAQTAEALQOSTNLTQMRDLISQANGSNDSEKTL-- 135
QY 115 HTTGKRGDLEIRVNSLQOVKSVGIDGIGETANGAFWVINITIRNDGSTPEVVD-GIFHL 173
DB 136 --NGEYKQLQ---KELDRISNTTTPGGRKLDGSEFGVASFGV--GSAANEIISVGIDEM 187
QY 174 --QNLNNGVOPDSTABIYANTNSGTIPDINPGVS--MTNLNVFDM 216
DB 188 SAESLNGTFKADGGCAVTAATASGTV--DLAIGITGSAAVNVAVDM 232

RESULT 13
AAU37120
ID AAU37120 standard; protein: 2344 AA.
XX
AC AAU37120;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1280.
XX
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX
KM antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
FN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
XX
PR 23-MAY-2000; 2000US-0206848P.
XX
PR 26-MAY-2000; 2000US-0207727P.
XX
PR 23-OCT-2000; 2000US-0242578P.
XX
PR 27-NOV-2000; 2000US-0253625P.
XX
PR 22-DEC-2000; 2000US-0257931P.
XX
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KU, Zyekind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX
DR N-PSDB; AAG54979.
XX
XX New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 12713; 511pp; English.

XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are *Bacteriella*
CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC [ftp.wipo.int/publ/published_pct_sequences](http://wipo.int/publ/published_pct_sequences)
XX
SQ Sequence 2344 AA;

Query Match 8.2%; Score 106; DB 4; Length 2344;
Best Local Similarity 25.9%; Pred. No. 9.5;
Matches 59; Conservative 18; Mismatches 79; Indels 72; Gaps 9;

QY 30 FAGSLMALIVFANTVVMIRGNTPRNASTVSATSLVNST-----NSQVAKQEQNSSTSPA 85
DB 89 FANSDAPLTSLELTQSTTVGN--QNSTTIDASTADSTVTKNSSV--QTSNSDVTSS 144
QY 86 HKS-----TNSLQHAQQAATTSQSQSLRYIPFHTTGKRGDLEIRVNSLQOVKSVGYD 139
DB 145 EKSENVSTSTNSNSQGEKLTSTSESTSSK-----NTTSSDPTKS--- 185
QY 140 GIGETANGAFWVINITIRNDGSTPEVVDGIFHLQNLNNGVOPDSTABIYANTNSGTIP 199
DB 186 -----TSTSTEQPI-----NTSTNQSTAS--NNTSOSTTP 214
QY 200 TDNLPGVSMNTNLVFDMPDFMTGHCYHSLVSMGFPQSDFTYVL 247
DB 215 TSANLMTKSTSTSTSTAPVKLTFSRL-----AMSTFASATTTAL 254

RESULT 14
ADJ34822
ID ADJ34822 standard; protein: 922 AA.
XX
AC ADJ34822;
XX
DT 22-APR-2004 (first entry)
XX
DE Xylanase from an environmental sample seq id 38.
XX
KW antibacterial; fungicide; thermostable xylanase activity;
XX
KW dough conditioning; beverage production; nutritional supplement;
XX
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;
XX
KW fungal infection; coccidiosis.
XX
OS Unidentified.
XX
FN WO2003106654-A2.
XX
PD 24-DEC-2003.
XX
PF 16-JUN-2003; 2003WO-US019153.
XX
PR 14-JUN-2002; 2002US-0389299P.
XX
PA (DIVE-) DIVERSA CORP.
XX
XX Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;
XX PI Steghehalian A;
XX

This Page Blank (uspto)

C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 05-Oct-2004
C/Accession: S53975
R/Contrib: R.; Churcher, C.M.
Submitted to the EMBL Data Library, April 1995
A/Reference number: S53969
A/Accession: S53975
A/Molecule type: DNA
A/Residues: 1-389 <CON>
A/Cross-references: UNIPROT:Q04951; UNIPARC:UPI00001356B9; EMBL:Z49212; NID:g798940; PIR
C/Genetics:
A/Gene: SGD:SCW10; MIPS:YMR305C
A/Cross-references: SGD:S0004921
A/Map position: 13R
C/Superfamily: family 17 glucosidase
C/Keywords: transmembrane protein
F/6-22/Domain: transmembrane #status predicted <TMM>

Query Match 8.6%; Score 111; DB 2; Length 389;
Best Local Similarity 23.1%; Pred. No. 0.44;
Matches 55; Conservative 40; Mismatches 95; Indels 48; Gaps 8;

QY 39 VPANTVYMRGNT-----PRNASTVSATTS-----LVNSTSSQVAKQDONS- 80
DB 36 VAAQVTVVVGSGNGETIVVNEAVVATTSSTAVASQATTSTLEPTTSANVTSSQOQTST 95
QY 81 -----STSPAHKSTNSLQHQHQQAATSSSQSK-LRIYIPHTYGVKVDLEI 125
DB 96 LOSSEAASTVSGSTSSSPSSSTSSASSASSASSASGALGITYSFPNDGSCCKSTQAQ 155
QY 126 RVNSLQOVYGVGDIGETANGAFWYINITIRNDGSTPMEVVDGIFHLQNLNGVYQDPS 185
DB 156 VASDLBQL--TGFDNRLXGVDSCQYENVLQAKTSQKFL--GIYVVKIQDAVDITIS 211
QY 186 TARIYANTNSGT--IPTDLNPGVSMTNLVFMDPDMTYGVGHYSLVAMGPFEGS 240
DB 212 AVESYGSWMDITTVSVGNELVNGSGSATTTQVGS-----YVSTKASLTSAGYVGS 261

RESULT 7
B84443
Hypothetical protein At2G03000 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: B84443
R/Lin: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: B84443
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-535 <STO>
A/Cross-references: UNIPROT:O80614; UNIPARC:UPI000004AC2A; GB:AE002093; NID:g3461824; PIR
C/Genetics:
A/Gene: At2G03000
A/Map position: 2

Query Match 7.9%; Score 102; DB 2; Length 535;
Best Local Similarity 24.3%; Pred. No. 3.3;
Matches 53; Conservative 31; Mismatches 88; Indels 46; Gaps 8;

QY 18 FLOGMNRILRPAGSLMALIVFANTVMIRGTPRNASTVSATTSVNSTSSQVAKOB 77
DB 277 YIQFGR-----LWE-----TNTT-----SSTRSRTTPTSTTNVPSNSSSRVLQTS 317
QY 78 QNS-STSPAHKSTN-SLQHQHQQAATSSSQSKLRIYIPHTYGVKVDLEIRVNSLQOVKS 135
DB 318 MSTRGTFPMSSSTRMSVQASMLAPGVSSMSSTVEIMPTSSARMTMSEVANSFQTSY 377
QY 136 VGVGDIGET-----ANG-----AFWVINTIRNDGSTPMEVVDGIF 171

DB 378 IQFGSLMNEFKVTXLSRVAVNENGEVCIQLEMRKEFTVSTSTNMRTSLMTLIRPID 437
QY 172 HLQNLNGVYQDPTARIYANTNSGTI--PTDLNPGVSM 208
DB 438 EQISLRATVSSSTRTRARLLTNESPATIRAVAMLPREVM 475

RESULT 8
F90073
Hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: F90073
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Chii, L.; Ogu-
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A69758; MUID:21311952; PMID:11418146
A/Accession: F90073
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-2271 <KUR>
A/Cross-references: UNIPROT:Q990Y4; UNIPARC:UPI00000CABB3; GB:BA000018; PIR:g13702612; F
A/Experimental source: strain N315
C/Genetics:
A/Gene: SA2447

Query Match 7.9%; Score 102; DB 2; Length 2271;
Best Local Similarity 25.0%; Pred. No. 22;
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSLMALIVFANTVMIRGTPRNASTVSATTSVNST---NSQVAKQDONSSTSPA 85
DB 89 FAASDAPLTSRLTQSEKTVGN--QNSTIEASTSTADSTVTKNSSSV--QTSNSDPTVSS 144
QY 86 HKS-----TNSLQHQHQQAATSSSQSKLRIYIPHTYGVKVDLEIRVNSLQOVYGVGD 139
DB 145 EKSEKVTSTVNSTSQOERLTSTSESTSSK-----NTTSSSDTYSVA-- 186
QY 140 GIGETANGAFWYINITIRNDGSTPMEVVDGIFHLQNLNGVYQDPTARIYANTNSGTIP 199
DB 187 -----STSTREPT-----NSTNSTSAS--NTSQSTIP 214
QY 200 TDLNPGVSMTNLVFMDPDMTYGVGHYSLVAMGPFEGSDETVAL 247
DB 215 SSVNLNKTSTSTSPAPVKLRTPSRL-----AMSTFASATTTAV 254

RESULT 9
A37853
Flagellin, 40K - Pseudomonas aeruginosa
C/Species: Pseudomonas aeruginosa
C/Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 26-Aug-1999
C/Accession: A37853
R/Totten, P.A.; Lory, S.
J. Bacteriol. 172, 7188-7199, 1990
A/Title: Characterization of the type a flagellin gene from Pseudomonas aeruginosa PAK.
A/Reference number: A37853; MUID:91072275; PMID:2123866
A/Accession: A37853
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-394 <TOT>
A/Cross-references: UNIPARC:UPI000016FCAL; GB:M57501; NID:g151225; PIR:AAA63458.1; PIR:
C/Superfamily: flagellin

Query Match 7.9%; Score 101.5; DB 2; Length 394;
Best Local Similarity 24.7%; Pred. No. 2.4;
Matches 55; Conservative 33; Mismatches 78; Indels 57; Gaps 11;

QY 36 ALIVFANTVMIRGTPRNASTVSA-----TTSLVNS-----T 68
DB 2 ALTVNTNLASL--NTQNLNNSASLNTSLQRLSTGSRINSAKDADAGLOIANRLTSGV 58

```

Qy 69 NSSQVAQOEONSSNSPA-----HKSTNSIQHQA---QAATSSSGSKARYPFHTYg 118
Db 59 NGLNVAATKNADGISLAQTAAGALQOOSTNIIQJRMHDLJLOSANGSNSDERTALN----- 113
Qy 119 KVGDLEIRVNSLQOVKSVGYDIGBTANGAFMVINITIRNGSTPMMEVVD-GIFHL-QN 175
Db 114 --GEKQLOKELDRISNTTTFGGRLDGSFCVASFOV--GSAANEIIISVGDIMSAS 168
Qy 176 LINGNVYQPDSTAEIYANTNSGTFPTDLNPGVS--MTTLVFPDM 216
Db 169 LNGTYFKADGGCAVTAATASGTV--DIAIGTGSASVANVKXDM 209

```

RESULT 10
E97130
uncharacterized secreted protein, homolog YXCC Bacillus subtilis [imported] - Clostridium
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: E97130
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: E97130
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-202 <KUR>
A/Cross-references: UNIPROT:Q97HY8; UNIPARC:UPI00000CA32C; GB:AE001437; PIDD:AAK79832.1
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC1868

[illegible]

```

RESULT 11
T13653
Hypothetical protein 95B7.2 - fruit fly (Drosophila melanogaster)
C|Species: Drosophila melanogaster
C|Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C|Accession: T13653
R|Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library April 1999
A|Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A|Reference number: Z17694
A|Accession: T13653
A|Status: Preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-982 <FER>
A|Cross-references: UNIPARC:UPI0000128461; EMBL:AL021728; NID:e1355643; PID:e1301389; P
C|Genetics;
A|Cross-references: FlyBase:FBgn0000376
A|Intons: 181/1; 249/1; 774/3; 899/1; 949/3
A|Note: EG:95B7.2

```

Query Match	7.64;	Score 98;	DB 2;	Length 982;
Best Local Similarity	20.34;	Pred. No. 15;		

	Matches	44;	Conservative	44;	Mismatches	89;	Indels	40;	Gaps	8;
Qy	17	LEIQGRNRI-IRFGSLMALI----	VFANTYMIKGNFRNASTVGAATISLVNSTNS	71						
Db	596	LYVFEQDRLBLEY---YWCVDSS	STAMASSILTIINVTSSSSVGOTGSSAGNAGT	652						
Qy	72	OYAKQKNSSTSPAHKSTNSLQAHQ-----	QAATSSSOSLRYIPF-----	114						
Db	653	SASSNCTYVASSNSSTSSSTSLPYNPNDCDS	VERARATYTTSSVSNLKHILLVANSLSRV	712						
Qy	115	-----HTYGVGLLEIRVNSLQQVKS	VGVDGIGETANGAFMWINITIRNDGSTPMEV	167						
Db	713	KRQCNCGHCTDRRDLMTKAQOLA	EATATG--GVGAWEGVGFRTPLFVR-----	PISNI	766					
Qy	168	DGIFHLQNLNGNYQPDSTAE--	IYATNNGSTPTDL	202						
Db	767	DPVRQLSALTRQKISRQVLVQRRL	LPTSSADRPYDL	803						

RESULT 12
 AD2672
 conserved hypothetical protein Atu0778 [imported] - Agrobacterium tumefaciens (strain C58)
 C|Species: Agrobacterium tumefaciens
 C|Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C|Accession: AD2672
 R|Wood, D.W.; Sebubal, J.C.; Kaul, R.; Monks, J.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.
 erge, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClellan,
 i Karo, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A|Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm,
 ster, R.W.
 A|Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A|Reference number: AB25777, MIMD:21608550, PMID:11743193
 A|Accession: AD2672
 A|Status: preliminary
 A|Molecule type: DNA
 A|Residues: 1-846 <KUR>
 A|Cross-references: UNIPROT:Q6UHA5, UNIPARC:UPI0000164523, GB:AB008668, PIDN:PAL41794.1,
 A|Experimental source: strain C58 (Dupont)
 C|Genetics:
 A|Gene: Atu0778
 A|Map position: circular chromosome

```

Query Match      7.54; Score 97; DB 2; Length 846;
Best Local Similarity 21.44; Pred. No. 14;
Matches    62; Conservative   30; Mismatches 100; Indels   98; Gaps   12;

QY      48 KGNTPPNASTVSAATTSLVNSTNSQVAKQ-----EONSSTSPAHKSTNLSLQAHQH 98
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      458 RGDIDPNTNRATYSTTSPASGCTSSPSSTSSASAPSGSSNYSTSTPATTTSCAASTPTAS 517
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      99 AATTSSSOSTLYR-----IPEHTYGKVGDELEIRVN-----                  128
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      518 SVSPSGSAATGHGISILPARSGIIDLDQY--AGTLIAARNACVGLINIANIVLDGAHGMKPR 575
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      129 -----SLQQYKSVGDIIGETANGAFWVINIT--KNDGSTPMEEVVDGIFFHLQNIN 177
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      576 SLLEPGRGQSFSISVDVGHND-GRDAGAFGIADIIGYGGLGGCAARAIARFGLYDORDIN 634
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      178 -----GNYQOPDSTAIEIVANTNSGTI-----PTDLNP 204
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      635 TCGDDFHHSGYIIAPEISLP.LPLDGLXATIGCYTAPORMSEIERGYLNGCMTDYSRGETDLN- 693
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      205 GVSMSTNLNVFMDMPFMYTG-----HYGOHYSILVASMGPFPGSDFTTYALP 248
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      694 --VMAAKLRDMWDALTI GENMLTPRYAGVTYA--KAEMAYYA--ETGGAPR 738
          : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 13
B97454
hypothetical protein AGR_C_1421 [imported] - Agrobacterium tumefaciens (strain C58, Cer
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C/Accession: B97454
 R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Hink, C.; Allinger, M.; Doughty, D.; Scott, C.; Leppas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A/Reference number: A97359; WUID:21608551; PMID:11743194
 A/Accession: B97454
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-848 <KUR>
 A/Cross-references: UNIPROT:Q8UHA5; UNIPARC:UPI00000D1910; GB:AB007869; PIDN:AAK86587.1; C/Genetics:
 A/Map position: circular chromosome

Query Match 7.5%; Score 97; DB 2; Length 848;
 Best Local Similarity 21.4%; Pred. No. 15;
 Matches 62; Conservative 30; Mismatches 100; Indels 98; Gaps 12;

48 RGNTPNASTVATTSVLNSTSSQVAKQ-----EONSTSPAHKSTNSLQHAHQ 98
 DB 460 RGDPTNRAVSTTPSASSGTTSSSPSSASPSGSSSNTSTPATTTGAASTPTAS 519
 QY 99 AATTSSSGSKLR-----IPFTYGVKVDLEIRVN----- 128
 DB 520 SVSSPSAAKTGHSILTPASGILDDQY--AGTIARFPAAGVGLNINIVLDGAKGMPK 577
 QY 129 -----SLQOVKSVGYDQIGETANGAFVWINITI--RNDSTPEWVDGIFHLQNLN 177
 DB 578 SLLEPGRQFSIVSDVGHND-GRDAAGAFGIADIGVFGDGAATRIAFGLYDGRDIN 636
 QY 178 -----GNVYQPDSTAEIYANTNSGTT-----PTDLP 204
 DB 637 TCGDPIHSGFYIAPRISLPLTDGLVATIGGYAPGMSIERGYLNGTMDYSRGETDLN- 695
 QY 205 GVSMTNLVFDMPDFTYTG-----HYGQHSILVAMSGFSGDETYPALP 248
 DB 696 --VMAKLFMDLALITGEMULTPYAGVTA-KAEMGAYA--ETGAPF 740

RESULT 14

T40634
 hypothetical protein SPBC685.03 - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: T40634
 R/Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, March 1999
 A/Reference number: Z21928
 A/Accession: T40634
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-452 <SEE>
 A/Cross-references: UNIPROT:Q9Y715; UNIPARC:UPI000006A6FA; EMBL:AL049474; PIDN:CA939360.
 A/Experimental source: strain 97zh-; cosmid c685
 C/Genetics:
 A/Map position: 2

Query Match 7.5%; Score 96.5; DB 2; Length 452;
 Best Local Similarity 23.2%; Pred. No. 6.9;
 Matches 54; Conservative 36; Mismatches 92; Indels 51; Gaps 8;

39 VFANTVYMRIGNTPRASVTSATTSLVNS-----TNSSQVAKQONSTSPAHK 87
 DB 114 VSSSLANSTYGTASNTQSSNSTSTMASTGVSAGSVPTNSTTMSISMSNSTSAATN 173
 QY 88 STNSLQHAQQAATTSSQSKLRYIPFATYGVKVDLEIRVN-SLQOVKSVG-----YDGI 141
 DB 174 TSSS-----SSSSSSQSVSVNSRIFSYFGLSQOYVNTSTRLCVGTPRANSTV 224
 QY 142 GTFANGAFVINITINDSTPM-----EVDGIFHLQNLNGVYQPDSTAEIYANTNS 195

DB 225 SVTNNGSA-VSNVTYNTNGMTSSNPKCYDDVVANI FGLDFTYAAVLSEVSLRSFALCNA 283
 QY 136 GTTPDLNPGVSMNTNLVFDMPFTYTG-----HYGQHSILVAMSGFSGSD 241
 DB 284 TT-----SSSLFRQIASVGVGSFHPSSSESGSPANLIGTNNYFMTD 325

RESULT 15

S19377
 probable membrane protein YCL048w - yeast (Saccharomyces cerevisiae)
 C/Species: Saccharomyces cerevisiae
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Oct-2004
 C/Accession: S19377
 R/Grenson, M.; Jauniaux, J.C.; Urrutazu, L.A. submitted to the Protein Sequence Database, March 1992
 A/Reference number: S19376
 A/Accession: S19377
 A/Molecule type: DNA
 A/Residues: 1-463 <GRS>
 A/Cross-references: UNIPROT:P25380; UNIPARC:UPI000013AF5C; EMBL:X59720; NID:G1907116; P1
 A/Map position: 3L
 A/Cross-references: SGD:S000053
 A/Map position: 3L
 C/Superfamily: sporulation-specific protein 2 precursor
 C/Keywords: transmembrane protein
 F/7-29/Domain: transmembrane #status predicted <TM1>
 F/446-463/Domain: transmembrane #status predicted <TM2>

Query Match 7.4%; Score 96; DB 2; Length 463;
 Best Local Similarity 22.2%; Pred. No. 7.8;
 Matches 51; Conservative 33; Mismatches 80; Indels 66; Gaps 9;

16 MLFIQGMNRI-----LRFASD-WALIVPANTVYMRIGNTPRNA 55
 DB 130 LIRIQAGNLNKEGLFQGLSLSVSRIPFLKFCGLSEKVVPLINLVSMDSQNIIRIK 189
 QY 56 STVSATTSVLNSTSSQVAKQONSTSPAHKSTNSLQHAQQAATTSSQSKLRYIPF 115
 DB 190 DIVISDTSLANENFNKV--QSIDT-----FNNNRPLETTHSNVKTIRGQ-----FS 236
 QY 116 TYGKVDLEIRVNSLQOVKSVGYDQIGETANGAFVWINITIRNDG-----STPE 165
 DB 237 VHAANKKLELMPHLRVE-----NITTRDSLVYLPOLTKVKSLE 278
 QY 166 VVDGIFHLQNLNGVYQPDSTAEIYANTNSGTT-----TDLPVGSMTTN 211
 DB 279 FLENFVELANL-NLQKIGTGLIINNVNLKVNLENLTDIOGLMIDN 327

Search completed: April 7, 2006, 16:24:07
 Job time : 30.2527 secs

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspio)


```
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=SA51720;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagsis K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571857; CAG43524.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 331 AA; 35855 MW; 753AF5FDAF50CF9 CRC64;

Query Match
Best Local Similarity 10.8%; Score 139; DB 2; Length 331;
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;

Qy 115 HTYG---KVGDLERVNSLQOVKSVGYGIGETANGAFWVINITIRNDGSTPMVEVDGIF 171
Db 81 HKIGETVNGDLEVTNVSVETKSVGPSIAPTNAGTIVADVITIKNGKALITIDSMF 140

Qy 172 HLOINANGVYOPDSTARIYANTN-SGTIPTD-----LNPGVSMNTNLVFDMPD 218
Db 141 KLLKS-GDKTFEADNTGSMANQNDNGSIENSFFLQRIINPDSTAQKIVFDVSE 192

RESULT 3
08NVZ1 STAAW PRELIMINARY; PRT; 331 AA.
AC 08NVZ1
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein MW1738.
GN OrderedLocusNames=MM1738;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=156620;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hizamatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL: BA000033; BAB95603.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 331 AA; 35855 MW; 753AF5FDAF50CF9 CRC64;

Query Match
Best Local Similarity 10.8%; Score 139; DB 2; Length 331;
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;

Qy 115 HTYG---KVGDLERVNSLQOVKSVGYGIGETANGAFWVINITIRNDGSTPMVEVDGIF 171
Db 81 HKIGETVNGDLEVTNVSVETKSVGPSIAPTNAGTIVADVITIKNGKALITIDSMF 140

Qy 172 HLOINANGVYOPDSTARIYANTN-SGTIPTD-----LNPGVSMNTNLVFDMPD 218
Db 141 KLLKS-GDKTFEADNTGSMANQNDNGSIENSFFLQRIINPDSTAQKIVFDVSE 192
```

```
Db 141 KLLKS-GDKTFEADNTGSMANQNDNGSIENSFFLQRIINPDSTAQKIVFDVSE 192

RESULT 4
06GF07 STAAW PRELIMINARY; PRT; 331 AA.
AC 06GF07
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=SA51880;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagsis K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571856; CAG40870.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 331 AA; 35850 MW; 82E9C361D2D59066 CRC64;

Query Match
Best Local Similarity 10.5%; Score 135; DB 2; Length 331;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

Qy 115 HTYG---KVGDLERVNSLQOVKSVGYGIGETANGAFWVINITIRNDGSTPMVEVDGIF 171
Db 81 HKIGETVNGDLEVTNVSVETKSVGPSIAPTNAGTIVADVITIKNGKALITIDSMF 140

Qy 172 HLOINANGVYOPDSTARIYAN-TNSGTIPTD-----LNPGVSMNTNLVFDMPD 218
Db 141 KLLKS-GDKTFEADNTGSMANQNDNGSIENSFFLQRIINPDSTAQKIVFDVSE 192

RESULT 5
05HEX9 STAAW PRELIMINARY; PRT; 331 AA.
AC 05HEX9
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Conserved domain protein, putative.
GN OrderedLocusNames=SA501847;
OS Staphylococcus aureus (strain COL).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=93062;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15774886; DOI=10.1126/JB.187.7.2426-2438.2005;
RA Gill S.R., Fouts D.E., Archer G.L., Mongodin E.F., DeBoy R.T.,
RA Ravel U., Paulsen I.T., Kolonay J.F., Brinkac L.M., Beanan M.J.,
RA Dodson R.J., Daugherty S.C., Madupu R., Angiuoli S.V., Durkin A.S.,
RA Hatt D.H., Vamathevan J.J., Khouli H., Uitterlinden T.R., Lee C.,
RA Dimitrov G., Jiang L., Qin H., Weidman J., Tran K., Kang K.H.,
RA Hance I.R., Nelson K.E., Fraser C.M.;
RT "Insights on evolution of virulence and resistance from the complete
RT genome analysis of an early methicillin-resistant Staphylococcus
RT aureus strain and a biofilm-producing methicillin-resistant
RT Staphylococcus epidermidis strain.";
RL J. Bacteriol. 187:2426-2438(2005).
```



```
DR InterPro; IPR006530; YD.  
DR Pfam; PF07564; DUF1542; 8  
DR Pfam; PF07554; FIVAR; 44.
```

DR Pfam; PF01468; GA; 46.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 1.
KW Complete proteome.
SQ SEQUENCE 6713 AA; 722339 MW; AF6EDE226BE488 CRC64;

Query Match 8.8%; Score 113.5; DB 2; Length 6713;
Best Local Similarity 24.4%; Pred. No. 53;
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;
QY 42 NTVMIRNTPRNAST--VSATSLVNSTNSQVAKQONSTSTRAHSTNSLQAHQCA 99
DB 899 NONTYDADESKKNATYQAVTAABGLINKQTGNTSKAVDNLAVTKAKALNAENLR 958
QY 100 ATTSQSQKLYIPHTYGVGDLIRVNSLQOVKSVGVDGI---GETANGAFVYINTI 156
DB 959 NAKTSATNTINGLPULTLQKONLGHQVQANV--VGWNGYKDKGNTLNTMGALRTSI 1016
QY 157 RNDGSTPEVVDGIFHLQNLNGVYQPDSTAEIYANTSGTIPIDLPGVSTNTNLVFD 216
DB 1017 QNDNTT--KTSQNYLDASDSNKNY---NTA---VNNANGVINTNPN--NMDANAINDM 1066

RESULT 14
086265_PSEAE PRELIMINARY; PRT; 387 AA.
ID 086265_PSEAE
AC 086265;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pfagellin.
DS Name=fl1C;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
CX NCBI_TaxID=287;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE:97072180; PubMed=8914989; DOI=10.1016/0014-5793(96)01099-X;
RA Spangenberg C., Heuer T., Burger C., Tummeler B.;
RT "Genetic diversity of flagellins of Pseudomonas aeruginosa."
RL PERS Lett. 396:213-217(1996).
DR EMBL; U81146; AAC28556.1; -; Genomic_DNA.
DR HSBP; O67803; IORV.
DR GO; GO:0009420; C:flagellar filament (sensu Bacteria); IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005159; P:ciliary or flagellar motility; IEA.
DR InterPro; IPR001029; Pfagellin_C.
DR Pfam; PF00700; Pfagellin_C_1.
DR Pfam; PF00669; Pfagellin_N_1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Pfagellin_C_1.
KW Pfagellum.
SQ SEQUENCE 387 AA; 39506 MW; 32E862D751F6142D CRC64;

Query Match 8.7%; Score 111.5; DB 2; Length 387;
Best Local Similarity 25.8%; Pred. No. 1.8;
Matches 57; Conservative 30; Mismatches 81; Indels 53; Gaps 10;

QY 36 ALIVANTVYVIRGTPRNASTVSA-----TSLVNS-----T 68
DB 2 ALVTNTIASL---NTQRIINNSASLNTSLQRLSTGRINSKADAGLQIANRLTSQV 58
QY 69 NSSQVAKQONSTSPA-----HKSTNSLOAHQ---QAATSSSQSKLYIPHTYG 118
DB 59 NGLNATVKNANDGISLAQTASGALQOSTNITLQRMEDLSQASNGSNDSERTAL---NG 114
QY 119 KYGDLIRVNSLQOVKSVGVDIGETANGAFVYINTIRNDGSTPEVVD-GIFHL--QN 175
DB 115 EYKQIQ---KELDRISNTTTFGRKLDSFGVAFFQV---GSANEIISVGIIDMSARS 168
QY 176 LINGVYQPDSTAEIYANTSGTIPIDLPGVSTNTNLVFD 216

DB 169 LNGTYFKADGGAAVTAATASGTVDIALGITGSTVNVKVD 209

RESULT 15
SCW10_YEAST STANDARD; PRT; 389 AA.
ID SCW10_YEAST
AC Q04951;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable family 17 glucosidase SCW10 precursor (EC 3.2.1.-) (Soluble cell wall protein 10).
DB Name=SCW10; Ordered locus names=YM9305C; ORF names=YM9952.07C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_TaxID=4932;
[1]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=S288C / AB972;
RC MEDLINE=97313268; PubMed=9169872;
RX Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T., Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S., Jagels K., Lye G., Moule S., Odeli C., Pearson D., Rajandream M.A., Rice P., Skellton J., Walsh S.V., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:90-93(1997).
[2]

RN PROTEIN SEQUENCE OF 30-39, AND SUBCELLULAR LOCATION.
RP STRAIN=S288C / SEY6210;
RC MEDLINE=98422453; PubMed=9748433;
RA Cappellaro C., Misa V., Tanner W.;
RT "New potential cell wall glucanases of Saccharomyces cerevisiae and their involvement in mating."
RL J. Bacteriol. 180:5030-5037(1998).
CC -1- FUNCTION: Glucanases possibly play a role in cell expansion during growth, in cell-cell fusion during mating, and in spore release during sporulation (By similarity).
CC -1- SUBCELLULAR LOCATION: Cell wall.
CC -1- PTM: Glycosylated.
CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 17 family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; Z49212; CA89138.1; -; Genomic_DNA.
DR PIR; S53975; S53975.
DR Geronline; 142983; -.
DR Ensembl; YMR305C; Saccharomyces cerevisiae.
DR SGD; S000004921; SCW10.
DR GO; GO:0009277; C:cell wall (sensu Fungi); IEA.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0000747; P:conjugation with cellular fusion; IGI.
DR InterPro; IPR000490; Glyco_hydro_17.
DR Pfam; PF00332; Glyco_hydro_17_1.
DR PROSITE; PS00587; GLYCOSYL_HYDROL_F17_1.
KW Cell wall; Complete proteome; Direct protein sequencing; Glycoprotein; Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 18 Potential.
FT PROPEP 19 29 Probable family 17 glucosidase SCW10.
FT CHAIN 30 389 Ser-rich.
FT COMBINS 65 134 Nucleophile (By similarity).
FT ACT SITE 326 326 Proton donor (By similarity).
FT ACT SITE 380 380 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 279 279
FT CONFLICT 38 39 NO -> Q8 (in Ref. 2).
SQ SEQUENCE 389 AA; 40469 MW; 21F394BD41337DE1 CRC64;

Query Match 8.6%; Score 111; DB 1; Length 389;

This Page Blank (uspto)

Sequence 19725, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 19725
 LENGTH: 417
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-19725

Query Match 8.3%; Score 106.5; DB 2; Length 417;
 Best Local Similarity 25.6%; Pred. No. 0.032;
 Matches 58; Conservative 33; Mismatches 79; Indels 57; Gaps 12;

QY 32 GSLMALIVPANTVMIRGTPRNASTVSA-----TISLVNS----- 67
 DB 21 GIKMLTVTVNTIASL---NTGRNLNNSASLNTSLQRLSTGSRINSAKODAGLQIANRL 77
 QY 68 ---TSSQVAKQBNSTSPA-----HKSTNSLOHQAHO---QAATSSSQSKRLRIYIPF 114
 DB 78 TSOVGNLAVATKANDGISLAQTAGALQOOSTNIIQRMEDLSLQSGNSGSDSERTAL-- 135
 QY 115 HTYQAGDLIRVNSLQOVKSVGDIQGTANGAFVINITIRNDSTPMRYVD-GIFHL 173
 DB 136 --NGEVKQLQ---KELDRISNTTTPGKRLDLSGFGVASFOV---GSAANEIISVGDIM 187
 QY 174 --ONINGNVYQPDSTAETIYANTNSGTIPTDLNPGVS--MTTNLVFDM 216
 DB 188 SABSINGTYFKADGGAVTAATASGTV--DIAGITGSAVAVKVD 232

RESULT 3
 US-10-172-502-4
 Sequence 4, Application US/10172502
 Patent No. 6841154
 GENERAL INFORMATION:
 APPLICANT: POSTER, Timothy et al.
 TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
 FILE REFERENCE: P07263US01/BAS
 CURRENT FILING DATE: 2002-06-17
 PRIOR APPLICATION NUMBER: US 60/298,098
 PRIOR FILING DATE: 2001-06-15
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 4
 LENGTH: 2283
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-10-172-502-4

Query Match 7.9%; Score 102; DB 2; Length 2283;
 Best Local Similarity 25.0%; Pred. No. 1.2;
 Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSLMALIVPANTVMIRGTPRNASTVSAATSLVNST---NSSQVAKQBNSTSPA 85
 DB 101 FAASDAPLTSBLNTQSRIVGN--QNSTTEASTSTADSTSVKNSSV--QTSNSDPTVS 156
 QY 86 HKS-----TNSLOHQAHOQAATSSSQSKRLRIIPFTYKGVGDLIRVNSLQOVKSVGDI 139
 DB 157 EKSEKVTSTNSTNSQOEKRLTSTESTSK-----NTTSSSDTKSVA-- 198

QY 140 GIGETANGAFVINITIRNDSTPMRYVDGIFHLQNLINGNVYQPDSTAETIYANTNSGTIP 199
 DB 199 -----STSTEQPI-----NTSTNGSTAS--NTNSQSTTP 226
 QY 200 TDLNPGVSMITNLVFDMPDFTYGHVGHYSLVASMGFPGSDETTYAL 247
 DB 227 SSVNLKNTSTSTSTAPVKLRTPFRL-----AMSTFASAAATTTAV 266

RESULT 4
 US-10-172-502-18
 Sequence 18, Application US/10172502
 Patent No. 6841154
 GENERAL INFORMATION:
 APPLICANT: POSTER, Timothy et al.
 TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
 FILE REFERENCE: P07263US01/BAS
 CURRENT FILING DATE: 2002-06-17
 PRIOR APPLICATION NUMBER: US 60/298,098
 PRIOR FILING DATE: 2001-06-15
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 18
 LENGTH: 485
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-10-172-502-18

Query Match 7.7%; Score 99; DB 2; Length 485;
 Best Local Similarity 24.5%; Pred. No. 0.23;
 Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 42 NTVMIRGTPRNASTVSAATSLVNST---NSSQVAKQBNSTSPAHS-----TNS 91
 DB 11 NTQSETVGN--QNSTTEASTSTADSTSVKNSSV--QTSNSDPTVSSEKSEKVTSTNS 66
 QY 92 LOHAHQQAATSSSQSKRLRIIPFTYKGVGDLIRVNSLQOVKSVGDIQGTANGAFV 151
 DB 67 TSNQEKLTSTSESTSK-----NTTSSSDTKSVA----- 96
 QY 152 INITIRNDSTPMRYVDGIFHLQNLINGNVYQPDSTAETIYANTNSGTIPTDLNPGVSMITN 211
 DB 97 -----STSTEQPI-----NTSTNGSTAS--NTNSQSTTPSSVNLKNTSTTS 136
 QY 212 LVFDMPDFTYGHVGHYSLVASMGFPGSDETTYAL 247
 DB 137 TSTAPVKLRTPFRL-----AMSTFASAAATTTAV 164

RESULT 5
 US-09-248-796A-15273
 Sequence 15273, Application US/09248796A
 Patent No. 6747137
 GENERAL INFORMATION:
 APPLICANT: Keith Weinstein et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 FILE REFERENCE: 107196.132
 CURRENT FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: US 60/074,725
 PRIOR FILING DATE: 1998-02-13
 PRIOR APPLICATION NUMBER: US 60/096,409
 PRIOR FILING DATE: 1998-08-13
 NUMBER OF SEQ ID NOS: 28208
 SEQ ID NO 15273
 LENGTH: 994
 TYPE: PRT
 ORGANISM: Candida albicans
 US-09-248-796A-15273

Query Match 7.5%; Score 97; DB 2; Length 994;

Best Local Similarity 22.3%; Pred. No. 1.1;
Matches 55; Conservative 34; Mismatches 82; Indels 76; Gaps 9;

QY 42 NNVNIRKGTTPRNASTVSAATSLV--NSTNSQVAKQONS-----TSPAKST 89
DB 147 STINOLASTJNSTLTITPTSTLVPDFSTNSRVSSTNRNSSESSMWLEKSSQLLS 206
QY 90 NSLOHQAQOATTSQ-----SKAYIPHTYGTGKVDLEIRNLSIQKSVYVDI 141
DB 207 TSPFNSSSTASTTESSELSAATSDSLNHSSSSVELSSSEADSSSESVETGSS 266
QY 142 GETANGAFWVNIITIRNDSTPMEVVDGIFHLQNLNGVYOPDST--ARIYANTSGTI 198
DB 267 DETAS-----NNSGDLFKALDTNAPPTVARSF--I 295
QY 199 PFDLNPV-----SMTNLVPMDFMTYTG-HVGOHSLVASMGP----- 237
DB 236 PLTIAGVYDNNKPIGTNKFYTNLILGNQDFWVYPLPYGLYWSKTSYGFVQHNVSVD 355
QY 238 -RGSDET 243
DB 356 VFGSINT 362

RESULT 6

US-09-228-986-72
; Sequence 72, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 910
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-72

Query Match 7.5%; Score 96.5; DB 2; Length 910;

Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 58; Conservative 40; Mismatches 98; Indels 61; Gaps 15;

QY 10 ILVPIYMLF---IQGRNRLR--FAGSLMALIVFANTVMIGNTPRNASTVSATTS 63
DB 302 IALFYNQSLGIVPMELGKLTQLRPFQGN-----YFISGNTWR---CPILTA 346
QY 64 LVNSTNSQVAKQONS-----TSPAKSTNSLOH---AQHQAATTSQSKLRYPH 115
DB 347 LNSCSDQLQYDSENNFTQGLPFSIGHLS-KKLYHLDGSLNLAGIPIPAIGNLSLFTL 405
QY 116 TYGK--VGDLERINSLIQKSVGVYDGIGETANGAFWVNIITIRNDSTPMEVVD---- 168
DB 406 NIGRNVFTGSIPISSLIMLQKRLTYMD-----SNNL-----QGNIPMEIGQLKSL 450
QY 169 GIFHL--QNLNGVYOPDSTAI---YANTSGTIPTDINPGVSMTTN-LVPDMPDFMT 221
DB 451 GLVLSGNLSGKI--PDFVANLQQLRYLYLNHNLGSDINANLKCVCVALLDLDSYNTL 508
QY 222 YGHVQHYSLVASMGP 238
DB 509 SGHIFQELAGLANIAFY 525

RESULT 7

US-10-101-464A-72
; Sequence 72, Application US/10101464A
; Patent No. 676804

GENERAL INFORMATION:

APPLICANT: Strabala, Timothy
APPLICANT: Nieuwenhuizen, Nicolaas
APPLICANT: Higgins, Colleen M.
TITLE OF INVENTION: Compositions Isolated from Plant Cells
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
FILE REFERENCE: 11000.1020c2
CURRENT APPLICATION NUMBER: US/10/101,464A
CURRENT FILING DATE: 2002-03-18
PRIOR APPLICATION NUMBER: 09/704,302
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 09/228,986
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/162,866
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US00/00724
PRIOR FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 989
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 72
LENGTH: 910
TYPE: PRT
ORGANISM: Pinus radiata
US-10-101-464A-72

Query Match 7.5%; Score 96.5; DB 2; Length 910;

Best Local Similarity 22.6%; Pred. No. 1.1;
Matches 58; Conservative 40; Mismatches 98; Indels 61; Gaps 15;

QY 10 ILVPIYMLF---IQGRNRLR--FAGSLMALIVFANTVMIGNTPRNASTVSATTS 63
DB 302 IALFYNQSLGIVPMELGKLTQLRPFQGN-----YFISGNTWR---CPILTA 346
QY 64 LVNSTNSQVAKQONS-----TSPAKSTNSLOH---AQHQAATTSQSKLRYPH 115
DB 347 LNSCSDQLQYDSENNFTQGLPFSIGHLS-KKLYHLDGSLNLAGIPIPAIGNLSLFTL 405
QY 116 TYGK--VGDLERINSLIQKSVGVYDGIGETANGAFWVNIITIRNDSTPMEVVD---- 168
DB 406 NIGRNVFTGSIPISSLIMLQKRLTYMD-----SNNL-----QGNIPMEIGQLKSL 450
QY 169 GIFHL--QNLNGVYOPDSTAI---YANTSGTIPTDINPGVSMTTN-LVPDMPDFMT 221
DB 451 GLVLSGNLSGKI--PDFVANLQQLRYLYLNHNLGSDINANLKCVCVALLDLDSYNTL 508
QY 222 YGHVQHYSLVASMGP 238
DB 509 SGHIFQELAGLANIAFY 525

RESULT 8

US-08-883-534-6
; Sequence 6, Application US/08883534
; Patent No. 5846777
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

RESULT 10
 US-09-107-532A-6217
 ; Sequence 6217, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arindello, Pamela Denek

```

;
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6217:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 522 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8) LOCATION 1...522
; SEQUENCE DESCRIPTION: SEQ ID NO: 6217:
US-09-107-532A-6217

Query Match 7.3%; Score 94.5; DB 2; Length 522;
Best Local Similarity 20.2%; Pred. No. 0.75;
Matches 50; Conservative 41; Mismatches 96; Indels 61; Gaps 11;

QY 37 LIVPANTVYMTIRGNTPRNASTVSATSL-VNSTSSQVAKQKQNSSTSPAHKSTNSLQHA 95
DB 211 LILAKSTIYGYGTPOPKSSHINASHMTNSTDTHLQKTNEGGLADPKQKVNLTGKI 270
QY 96 Q-----HOAATSSSQSKLRYIPFHT-YGKVDELIRVNSLQOVKSVGY--DGIGET 144
DB 271 NDMLDITETHSLVMTTDSGAMVWVPFIDSQMLRBERYLTGLIRE---GYMDNTYDI 326
QY 145 ANGAFWVN-----ITIRNDGSTPMEVY-----DGIFHLQNLNG 178
DB 327 SADPSWFIYSSGDDVRLITSRDNGKTMQTSIIITGQVPIRYKAVFYGNQDLFICTASG 386
QY 179 NVYQPDSTAEIYANTNSGTIPDLNPGVSMTNLVFDPDPMTYGHVQGHYSLV-ASNGF 237
DB 387 ---MSAEFLTYITWDGI--TWIKSGSTMINQPV-----QNVSPVTSISGF 428

QY 238 FGSDETTY 245
DB 429 VSTRSLP 436

RESULT 11
US-09-009-620-2
; Sequence 2, Application US/09009620A
; Patent No. 6127603
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357C
; CURRENT APPLICATION NUMBER: US/09/009,620A
; CURRENT FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-009-620-2

Query Match 7.3%; Score 94; DB 2; Length 349;
Best Local Similarity 21.3%; Pred. No. 0.45;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;
```

```

QY 102 TSSQSKL-----RYIPFH-----TYGKVGDEIRVNSLQOVKSVGYDGIGETANGAF 149
DB 127 TSKAVNRLTANGKEFVANNNTAHPKTYVDKIVIPKIKGLKQPSLSQDDIALGN--- 183
QY 150 WVINI-----IRNDGSTPMEVVDIPIHLQNLNGVYQPDSTAEIYANT---NSGTI 198
DB 184 -VKNIKRVNGKYYTYKEDGTIQKN-----YALNINNGKTFPDETGAISNNTLPSKKGNI 236
QY 199 PTDLNPGVSMTNLVFDPDPMTYGHVQGHY 229
DB 237 TNNNDTNSFAQYNGVYS-TDVANPEHY-DHY 265

RESULT 12
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-4

Query Match 7.3%; Score 94; DB 2; Length 1375;
Best Local Similarity 21.3%; Pred. No. 3.7;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;

QY 51 TPRNASTVSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAHQHQAAT 101
DB 67 TATDTSTATSATISQPTATVTDVNSTNSNTTANTANFVVKPTTTSQAATDSDKIIT 126
QY 102 TSSQSKL-----RYIPFH-----TYGKVGDEIRVNSLQOVKSVGYDGIGETANGAF 149
DB 127 TSKAVNRLTANGKEFVANNNTAHPKTYVDKIVIPKIKGLKQPSLSQDDIALGN--- 183
QY 150 WVINI-----IRNDGSTPMEVVDIPIHLQNLNGVYQPDSTAEIYANT---NSGTI 198
DB 184 -VKNIKRVNGKYYTYKEDGTIQKN-----YALNINNGKTFPDETGAISNNTLPSKKGNI 236
QY 199 PTDLNPGVSMTNLVFDPDPMTYGHVQGHY 229
DB 237 TNNNDTNSFAQYNGVYS-TDVANPEHY-DHY 265

RESULT 13
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
```

FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 7.3%; Score 94; DB 2; Length 1375;
Best Local Similarity 21.3%; Pred. No. 3.7;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;

QY 51 TPRASVATSVATSLVNSTSSQVAKQONSTSTSPAHKSTNSLQHQ 101
DB 67 TATSTSTASATSPRTATVTDNSTQSTNTTANTANFVYKPTTSEQAKTDNSDKIT 126
QY 102 TSSQSKL-----RYIPFH-----TYGKVDLEIRVNSLQVYKVGVDIGETANGAF 149
DB 127 TSKAVNLTATGKVPANNNTAHPKTYTDKIVPKPKIGKLPSSLSQDDIALGN--- 183
QY 150 WVINI-----IRNDSTPEMEVVDGIFHLQNLNGVYQPDSTAETIYANT---NSGTI 198
DB 184 -VKNIRKVGKYYKKEDGTLQKN-----YALNINGTFFPDGTALSNNTLSKKGNI 236
QY 199 PTDINPGVSMTTNLVFPMDPFTYGVHGOHY 229
DB 237 TNNNDNTNSFAQYNQVYS-TDVANFEHV-DHY 265

RESULT 14
US-09-270-767-31859
Sequence 31859, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 31859
LENGTH: 313
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-31859

Query Match 7.3%; Score 93.5; DB 2; Length 313;
Best Local Similarity 25.7%; Pred. No. 0.43;
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;

QY 46 MIRGNTPRNASTVATSL-----VNSTSSQVAKQ---EONSTSPAHKSTNSLQHQ 94
DB 142 VSTNTSTLTNTTASTSTASIMHRTKTDISHNSSGAGSQFYCELNSTSTSRHNSLDRDVH 201
QY 95 AQHQA-ATTSSSQSKLRYIPFTYGVKVDLEIRVNSLQVYKVGVDIGETANGAFWVIN 153

DB 202 HQHASLMSASSSVSNLSLDSNSGGR-----QSGKLTTHSQSWG-NGLKSHGNGT----- 249
QY 154 ITRNDGSTPEMEVVDGIFHLQNLNGVYQPD 185
DB 250 -----TNGGSP-----HINAQLVOPSS 266

RESULT 15
US-09-270-767-47076
Sequence 47076, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 47076
LENGTH: 313
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-47076

Query Match 7.3%; Score 93.5; DB 2; Length 313;
Best Local Similarity 25.7%; Pred. No. 0.43;
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;

QY 46 MIRGNTPRNASTVATSL-----VNSTSSQVAKQ---EONSTSPAHKSTNSLQHQ 94
DB 142 VSTNTSTLTNTTASTSTASIMHRTKTDISHNSSGAGSQFYCELNSTSTSRHNSLDRDVH 201
QY 95 AQHQA-ATTSSSQSKLRYIPFTYGVKVDLEIRVNSLQVYKVGVDIGETANGAFWVIN 153
DB 202 HQHASLMSASSSVSNLSLDSNSGGR-----QSGKLTTHSQSWG-NGLKSHGNGT----- 249
QY 154 ITRNDGSTPEMEVVDGIFHLQNLNGVYQPD 185
DB 250 -----TNGGSP-----HINAQLVOPSS 266

Search completed: Apr 11 7, 2006, 16:25:19
Job time : 37.5187 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:40:31 ; Search time 121.002 Seconds

(without alignments)
856.362 Million cell updates/sec

Title: US-10-784-592-43

Perfect score: 1289
Sequence: 1 KRIMVGLMIVPYIMLFIO.....YSLVSMGPRFGSDETTYALP 248

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA Main:

1: /cgn2_6/pcodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/pcodata/1/pubppaa/US08_PUBCOMB.pep.*
3: /cgn2_6/pcodata/1/pubppaa/US09_PUBCOMB.pep.*
4: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep.*
5: /cgn2_6/pcodata/1/pubppaa/US10_PUBCOMB.pep.*
6: /cgn2_6/pcodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	100.0	248	US-10-784-592-43	Sequence 43, Appl
2	135	10.5	331	US-09-815-242-5912	Sequence 5912, Ap
3	135	10.5	331	US-09-815-242-12872	Sequence 12872, A
4	135	10.5	331	US-09-815-242-13146	Sequence 13146, A
5	135	10.5	331	US-10-282-122A-44046	Sequence 44046, A
6	113.5	8.8	6713	US-10-282-122A-43811	Sequence 43811, A
7	107.5	8.3	5795	US-09-815-242-5639	Sequence 5639, Ap
8	107.5	8.3	5795	US-09-815-242-12610	Sequence 12610, A
9	106	8.2	2344	US-09-815-242-12713	Sequence 12713, A
10	105.5	8.2	1020	US-10-437-963-131231	Sequence 131231, A
11	105	8.1	2386	US-10-156-761-7751	Sequence 7751, Ap
12	102	7.9	2261	US-10-470-048B-60	Sequence 60, Appl
13	102	7.9	2271	US-10-282-122A-43924	Sequence 43924, A
14	102	7.9	2383	US-10-172-502-4	Sequence 4, Appl1
15	102	7.9	2383	US-11-020-509-4	Sequence 4, Appl1
16	101.5	7.9	394	US-10-125-692-14	Sequence 14, Appl1
17	101.5	7.9	394	US-10-991-347-14	Sequence 14, Appl1
18	100	7.8	982	US-11-097-143-4851	Sequence 4851, Ap
19	100	7.8	982	US-11-097-143-28995	Sequence 28995, A
20	99	7.7	194	US-10-282-122A-57709	Sequence 57709, A
21	99	7.7	485	US-10-172-502-18	Sequence 18, Appl
22	99	7.7	485	US-11-020-509-18	Sequence 18, Appl
23	98.5	7.6	2398	US-10-282-122A-70176	Sequence 70176, A
24	98.5	7.6	6281	US-09-815-242-13996	Sequence 13996, A
25	98.5	7.6	10498	US-10-470-048B-440	Sequence 440, App
26	97.5	7.6	393	US-10-437-963-127450	Sequence 127450, A
27	96.5	7.5	910	US-10-101-464A-72	Sequence 72, Appl

28	96.5	7.5	910	US-10-864-252-72	Sequence 72, Appl
29	96	7.4	2233	US-10-998-197-16	Sequence 16, Appl
30	95.5	7.4	632	US-10-040-906A-4	Sequence 4, Appl1
31	95.5	7.4	632	US-11-098-545-4	Sequence 4, Appl1
32	95.5	7.4	786	US-10-369-493-16335	Sequence 16335, A
33	95.5	7.4	791	US-10-369-493-15596	Sequence 15596, A
34	95.5	7.4	805	US-10-369-493-15973	Sequence 15973, A
35	94.5	7.3	458	US-10-282-122A-53974	Sequence 53974, A
36	94.5	7.3	627	US-10-040-906A-6	Sequence 6, Appl1
37	94.5	7.3	627	US-11-098-545-6	Sequence 6, Appl1
38	94.5	7.3	632	US-11-067-557-2	Sequence 2, Appl1
39	94.5	7.3	633	US-11-067-557-32	Sequence 32, Appl
40	94.5	7.3	633	US-11-067-557-38	Sequence 38, Appl
41	94.5	7.3	634	US-11-067-557-166	Sequence 166, App
42	94.5	7.3	634	US-11-067-557-224	Sequence 224, App
43	94	7.3	1375	US-09-740-274-4	Sequence 4, Appl1
44	94	7.3	1375	US-10-383-930-35	Sequence 35, Appl
45	94	7.3	1375	US-10-797-821-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1					
US-10-784-592-43					
Sequence 43, Application US/10784592					
Publication No. US20050147983A1					
GENERAL INFORMATION:					
APPLICANT: Willing, Reinhard					
APPLICANT: Oestergaard, Peter					
APPLICANT: Lassen, Soren					
TITLE OF INVENTION: POLYPEPTIDES OF ALICYCLOBACILLUS SP.					
FILE REFERENCE: 10406.203-US					
CURRENT APPLICATION NUMBER: US/10/784,592					
CURRENT FILING DATE: 2004-02-23					
NUMBER OF SEQ ID NOS: 55					
SOFTWARE: PatentIn version 3.2					
SEQ ID NO 43					
LENGTH: 248					
TYPE: PRT					
ORGANISM: Alicyclobacillus sp.					
FEATURE:					
NAME/KEY: SIGNAL					
LOCATION: (1)..(41)					
FEATURE:					
NAME/KEY: mat.peptide					
LOCATION: (42)..(248)					
OTHER INFORMATION: functional polypeptide					
US-10-784-592-43					
Query Match					
Best Local Similarity 100.0%, Pred. No. 3.2e-11;					
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	KRIMVGLMIVPYIMLFIOGRNRILRFAGSLALIVPANTVMIRKNTPRNSTVSA	60		
DB	1	KRIMVGLMIVPYIMLFIOGRNRILRFAGSLALIVPANTVMIRKNTPRNSTVSA	60		
QY	61	TTSLVNSTNSGOVAQOBNSSSTSPAHKSTNSLOHQAATSSSOSKLRYPPTHYGV	120		
DB	61	TTSLVNSTNSGOVAQOBNSSSTSPAHKSTNSLOHQAATSSSOSKLRYPPTHYGV	120		
QY	121	GDLEIRVNSLOQVKSVDGIGETANGAFVWVITIRNDGSTPMVEVDGIFHLQNLNGV	180		
DB	121	GDLEIRVNSLOQVKSVDGIGETANGAFVWVITIRNDGSTPMVEVDGIFHLQNLNGV	180		
QY	181	YQPDSTAIRVANTNSGTLPTDLNPGVSMVTNLVFMDFPMYTGAVGHVSLVSMGPRFGS	240		
DB	181	YQPDSTAIRVANTNSGTLPTDLNPGVSMVTNLVFMDFPMYTGAVGHVSLVSMGPRFGS	240		
QY	241	DETTYALP 248			
DB	241	DETTYALP 248			

RESULT 2

```
US-09-815-242-5912
; Sequence 5912, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5912
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5912
```

```
Query Match      10.5%; Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY      115 HTYG---KVGDLIRVNSLQGVKSYGDIETANGAFVYVINITRNDGSTPMEVVDGIF 171
      81 HRIGETVKNGLDEVTVNSVETKSVGPSLAPTNAGIFVADVITIKNGKEALTTIDSSMF 140
      172 HLOINGNVYOPDSTAETIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
      141 KLSK-GDKTFEADNTGSMNSQSDNGSIENSFFLQRIINPDSTAQCKIVFDVSE 192
DB
```

RESULT 3

```
US-09-815-242-12872
; Sequence 12872, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12872
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12872
```

```
Query Match      10.5%; Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY      115 HTYG---KVGDLIRVNSLQGVKSYGDIETANGAFVYVINITRNDGSTPMEVVDGIF 171
      81 HRIGETVKNGLDEVTVNSVETKSVGPSLAPTNAGIFVADVITIKNGKEALTTIDSSMF 140
      172 HLOINGNVYOPDSTAETIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
      141 KLSK-GDKTFEADNTGSMNSQSDNGSIENSFFLQRIINPDSTAQCKIVFDVSE 192
DB
```

RESULT 4

```
US-09-815-242-13146
; Sequence 13146, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13146
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
```

US-09-815-242-13146

Query Match 10.5%; Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 115 HTYG---KVGDEIRVNSIQOVKSVGYDGIGETANGAFVINITIRNDGSTMEVVDGIF 171
DB 81 HXIGETVKNKGDLEIVTVNSVETKMSVGPSPALPTNAKGIFFVADVITIKKKEKALTTIDSSMF 140

QY 172 HLQNLNGVYQPDSTABIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
DB 141 KUKS-GDKTFEADNTGSMNSQDNGSINSFPLQINPDSIAQGIYFVDSB 192

RESULT 5
US-10-282-122A-44046
Sequence 44046, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR FILING DATE: 2003-02-20

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2001-02-09

PRIOR FILING DATE: 2001-02-09

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

us-10-784-592-43.rapbm

QY 172 HLQNLNGVYQPDSTABIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
DB 141 KUKS-GDKTFEADNTGSMNSQDNGSINSFPLQINPDSIAQGIYFVDSB 192

RESULT 6
US-10-282-122A-43811
Sequence 43811, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR FILING DATE: 2003-02-20

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2001-02-09

PRIOR FILING DATE: 2001-02-09

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

us-10-784-592-43.rapbm

QY 172 HLQNLNGVYQPDSTABIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
DB 141 KUKS-GDKTFEADNTGSMNSQDNGSINSFPLQINPDSIAQGIYFVDSB 192

RESULT 7
US-10-282-122A-43811
Sequence 43811, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A

CURRENT FILING DATE: 2003-02-20

PRIOR FILING DATE: 2003-02-20

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-03-21

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-23

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-06

PRIOR FILING DATE: 2000-09-09

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-10-23

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-11-27

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2001-02-09

PRIOR FILING DATE: 2001-02-09

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

PRIOR FILING DATE: 2001-02-16

```
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 5639
; LENGTH: 2086
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5639

Query Match      8.3%; Score 107.5; DB 3; Length 2086;
Best Local Similarity 24.2%; Pred. No. 3.4;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

QY 42 NTVTWIRGNTPRNAST--VSATTSILVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQHOA 99
DB 1323 NQNTYDADBSKRNATQAVTAABGILNKQGTGTSKADVDNALNTVTAKAALNQAENLR 1382
QY 100 ATTSSQSKLRYIPHTYGVKGDLEIRVNSLQOVKSV-GYDGIGETANGAFVINITIRN 158
DB 1383 NTKTSATNTINGLPULTQKQDNLKHQVEQKQNVAGVGVKDKGNTLTNTAGALRTSICN 1442
QY 159 DQSTPEWVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTFPLDLPVGSMTTNLVFDM 216
DB 1443 DMTT--KTSQNYVLADSDSKNNY---NTA---VNNANGVINTVNP--NMDANALNGM 1490

RESULT 8
US-09-815-242-12610
; Sequence 12610, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12610
; LENGTH: 5795
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12610

Query Match      8.3%; Score 107.5; DB 3; Length 5795;
Best Local Similarity 24.2%; Pred. No. 14;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

QY 42 NTVTWIRGNTPRNAST--VSATTSILVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQHOA 99
DB 3600 NQNTYDADBSKRNATQAVTAABGILNKQGTGTSKADVDNALNTVTAKAALNQAENLR 3659
QY 100 ATTSSQSKLRYIPHTYGVKGDLEIRVNSLQOVKSV-GYDGIGETANGAFVINITIRN 158
DB 3660 NTKTSATNTINGLPULTQKQDNLKHQVEQKQNVAGVGVKDKGNTLTNTAGALRTSICN 3719
QY 159 DQSTPEWVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTFPLDLPVGSMTTNLVFDM 216
DB 3720 DMTT--KTSQNYVLADSDSKNNY---NTA---VNNANGVINTVNP--NMDANALNGM 3767

RESULT 9
US-09-815-242-12713
; Sequence 12713, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
```

SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12713
; LENGTH: 2344
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12713

Query Match 8.2%; Score 106; DB 3; Length 2344;
Best Local Similarity 25.9%; Pred. No. 5.5;
Matches 59; Conservative 18; Mismatches 79; Indels 72; Gaps 9;

QY 30 FAGSLMALIVFANTVMIRGNTPRNASTVSAATSLVNST---NSSQVAKOEONSTSPA 85
DB 89 FAASDAPLSEINTQSEITGN--QNSTTIDASTSTDSYTKNSSSV--QTSNEDTVSS 144
QY 86 HKS-----TNSLQHAQHOAATSSSQSLKRYIPFHTYKVGDLIRVNSLQOVKSVGYD 139
DB 145 EKSENVSTYTNSTNSQOBLTSTSEBTSK-----NTSSSDTKSV--- 185
QY 140 GIGETANGAFWYINTIRNDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTFP 199
DB 186 -----TSTSTTEOP-----NTSTNOSTAS--NNTSQSTTP 214
QY 200 TDLNPGVSWTNLVFEDMPFTYGHVGHYSLVSWGFPSPDETTVAL 247
DB 215 TSNANKTSTSTSTRAPVCLRTFSRL-----AMSTASAAITTL 254

RESULT 10
US-10-437-963-131231

; Sequence 131231, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuc, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 131231

; LENGTH: 1020

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_33316C.1.pcp

US-10-437-963-131231

Query Match 8.2%; Score 105.5; DB 4; Length 1020;
Best Local Similarity 25.1%; Pred. No. 1.9;
Matches 51; Conservative 32; Mismatches 63; Indels 57; Gaps 11;

QY 42 NTVVMIRGTPRNASTVSAATSLVNSTSSQVAKOEONSTSPAKSTNSLQHAQHOAT 101
DB 573 NTVFISGGK-----SASVSLMSTLHSSQSE-----NSNMQIGQ----- 607
QY 102 TSSSQSLKRYIPFHTYKVGDLIRVNSLQOVKSV-----VGYDQIGR--TANGAFW 150
DB 608 ---GGLKQ-----YNAAGALSLVNSKTYEAPNEQKKYKYYIGQDNGSGSPQASGAYI 658
QY 151 VT---NITRNDGSTPMEVVDGIF---HLQNLNGVYQPDST--AEIYANTN--SGTIP 199
DB 659 FRPNGTVPKTDGOVPLTYLRGSIIDEVH--QOINPMIYQINRVYKGVKQVETETFPVGP 717
QY 200 TDLNPGVSWTNLVFEDMPFTY 222

DB 718 VDDGNGKELSTEVYTNMAFNKTF 740

RESULT 11
US-10-156-761-7751

; Sequence 7751, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIRA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 7751

; LENGTH: 2386

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-7751

Query Match 8.1%; Score 105; DB 4; Length 2386;
Best Local Similarity 22.2%; Pred. No. 7;
Matches 52; Conservative 30; Mismatches 92; Indels 60; Gaps 10;

QY 55 ASTVSATSLVNSTSSQVAKOEONSTSPAKSTNSLQHAQHOATSSSQSLKRYIPF 114
DB 1364 SKVVEHTSSATTTGYTAKGELAQITDPKGNATLYTDMAAQKTTDDPAGSSSEY 1423
QY 115 HTYGRVGDLEIRVNSLQOVKSVGYDQIGRTANG-----AFWVINTIRNDGSTPMEV 167
DB 1424 NENGVSQTTATNDVQTVLTYGVNLSRATSVRSAGDELAAMVDDPAATGKG--QIT 1481
QY 168 DGFHLQNLNGVYQPDSTAIYANTNSG-----TYTDLNPGVS--MTTNL 212
DB 1482 SAV--SRDASNTY-----TTKYGKEDRGRPLATVTLPTVN--GIAGDYTTSV 1528
QY 213 VEDMPDM-----TYGHVGHYSLVSWGFPSS---DETTY 245
DB 1529 TYDAADHTSVSYPAAGLAERKVTYYDYDQPTRLTSLG--GTAYIDNTTY 1580

RESULT 12
US-10-470-048B-60

; Sequence 60, Application US/10470048B

; Publication No. US2005003744A1

; GENERAL INFORMATION:

; APPLICANT: MEINKE ET AL.

; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF

; FILE REFERENCE: S0NN:03505

; CURRENT APPLICATION NUMBER: US/10/470,048B

; CURRENT FILING DATE: 2003-07-25

; NUMBER OF SEQ ID NOS: 603

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 60

; LENGTH: 2261

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

Query Match 7.9%; Score 102; DB 5; Length 2261;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSIMALIVFANTVYMRGNTPRNASTVASTSLVNST---NSSQVAKQONSSTSPA 85
DB 89 FAASDAPLSTSLANTQSEVGN--QNSTTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 144
QY 86 HKS-----TNSLQHAQHOAATTSSQSKLARIIPHTYGVKXGDLIRVNSLQOVKSVGYD 139
DB 145 EKSEKVTSTSTNSQOEKLTSTSESTSSK-----NTTSSSDTKSVA-- 186
QY 140 GIGETANGAFVINITINDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTIP 199
DB 187 -----STSTEQPI-----NTSTNGSTAS--NTTSQSSTTP 214
QY 200 TDLNPGVSMITNLVFDMPDFMTYGHVGHYSLVASMGEFGSDETTYAL 247
DB 215 SSVNLKSTSTSTSTAPVKLRTFSRL-----AMSTFASAAATTAV 254

RESULT 13

US-10-282-122A-43924
; Sequence 43924, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43924
; LENGTH: 2271
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43924

Query Match 7.9%; Score 102; DB 4; Length 2271;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;
30 FAGSIMALIVFANTVYMRGNTPRNASTVASTSLVNST---NSSQVAKQONSSTSPA 85

DB 89 FAASDAPLSTSLANTQSEVGN--QNSTTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 144
QY 86 HKS-----TNSLQHAQHOAATTSSQSKLARIIPHTYGVKXGDLIRVNSLQOVKSVGYD 139
DB 145 EKSEKVTSTSTNSQOEKLTSTSESTSSK-----NTTSSSDTKSVA-- 186
QY 140 GIGETANGAFVINITINDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTIP 199
DB 187 -----STSTEQPI-----NTSTNGSTAS--NTTSQSSTTP 214
QY 200 TDLNPGVSMITNLVFDMPDFMTYGHVGHYSLVASMGEFGSDETTYAL 247
DB 215 SSVNLKSTSTSTSTAPVKLRTFSRL-----AMSTFASAAATTAV 254

RESULT 14

US-10-172-502-4
; Sequence 4, Application US/10172502
; Publication No. US20030185833A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2283
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-4

Query Match 7.9%; Score 102; DB 4; Length 2283;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSIMALIVFANTVYMRGNTPRNASTVASTSLVNST---NSSQVAKQONSSTSPA 85
DB 101 FAASDAPLSTSLANTQSEVGN--QNSTTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 156
QY 86 HKS-----TNSLQHAQHOAATTSSQSKLARIIPHTYGVKXGDLIRVNSLQOVKSVGYD 139
DB 157 EKSEKVTSTSTNSQOEKLTSTSESTSSK-----NTTSSSDTKSVA-- 198
QY 140 GIGETANGAFVINITINDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTIP 199
DB 199 -----STSTEQPI-----NTSTNGSTAS--NTTSQSSTTP 226
QY 200 TDLNPGVSMITNLVFDMPDFMTYGHVGHYSLVASMGEFGSDETTYAL 247
DB 227 SSVNLKSTSTSTSTAPVKLRTFSRL-----AMSTFASAAATTAV 266

RESULT 15

US-11-020-509-4
; Sequence 4, Application US/11020509
; Publication No. US20050106648A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US02/BAS
; CURRENT APPLICATION NUMBER: US/11/020,509
; CURRENT FILING DATE: 2004-12-27
; PRIOR APPLICATION NUMBER: US 10/172,502
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4
; LENGTH: 2283
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-11-020-509-4

Query Match 7.9%; Score 102; DB 6; Length 2283;
Best Local Similarity 25.0%; Pred. No. 12;
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSLMALIVPANTVTMIRGNTPRNSTVSATTSLVNST---NSSQVAKQONSSSTSPA 85
DB 101 FAASDBAPLITSELTNGSETVGN--QNSTTIEASTADSTSVTKSSSV--QTSNSDPTVSS 156
QY 86 HKS-----TNSIQHQHQAATSSSQSLKRIIPHTYGVGDLERVNSLQOVKSYGYD 139
DB 157 EKSEKVTSTNSTNSQOEKLTSTSESTSK-----NTTSSSDPTKSYA-- 198
QY 140 GIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLNGNVYQPDSTAETIYANTNSGTIP 199
DB 199 -----STSTEQPI-----NTSTNOSTAS--NNTSQSTTP 226
QY 200 TDINPGVSMITTNLVFPMDPFTMYGHVGOHYSLVASMGFFGSDETTYAL 247
DB 227 SSVNLNKTSTTSTSTAPVKLRTPSRL-----AMSTFASAAATTAY 266

Search completed: April 7, 2006, 16:44:39
Job time : 123.002 secs

This Page Blank (uspto)

Db : : : :
262 SSSYSGAD 270

RESULT 5

US-11-096-568A-22862
; Sequence 22862, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 8750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22862
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1)..(333)
; OTHER INFORMATION: Ceres Seq. ID no. 12410414
US-11-096-568A-22862

Query Match 7.0%; Score 90.5; DB 7; Length 333;
Best Local Similarity 19.0%; Pred. No. 0.84;
Matches 36; Conservative 30; Mismatches 72; Indels 51; Gaps 7;

QY 56 STVSATTSLVNSTNSQVAKQONSTSPAHKSTNSLQAHQAATSSQSKLRITPFI 115
Db 168 SYMPMDTATVRNTNANAVPSPMDNTLAQPF-----SAGVQGAATNCCSSM----- 215
QY 116 TYGKVGDLEIRVNSLQOVKSVGYDGGTANGAFWVINITIRNDGST-PMEVVDGIFHLQ 174
Db 216 -----ESPSCG-WTSSAVRGEMNVPLRAMPDPAQVY 247
QY 175 NUNGNVQPDSTAETAIYANTNSGTITPDLNPGVSMITNLVFDL--PDMTYGHVGQHSIV 232
Db 248 NPLGSIPTDSTSGHL--QMLKAMDVPIETALLMKNLSMNLRSPDF-----EQHRRLL 299
QY 233 ASMGFFGSD 241
Db 300 SSSYSGAD 308

RESULT 6

US-11-096-568A-22861
; Sequence 22861, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22861
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: Ceres Seq. ID no. 12410413
US-11-096-568A-22861

Query Match 7.0%; Score 90.5; DB 7; Length 336;
Best Local Similarity 19.0%; Pred. No. 0.85;
Matches 36; Conservative 30; Mismatches 72; Indels 51; Gaps 7;

QY 56 STVSATTSLVNSTNSQVAKQONSTSPAHKSTNSLQAHQAATSSQSKLRITPFI 115
Db 171 SYMPMDTATVRNTNANAVPSPMDNTLAQPF-----SAGVQGAATNCCSSM----- 218
QY 116 TYGKVGDLEIRVNSLQOVKSVGYDGGTANGAFWVINITIRNDGST-PMEVVDGIFHLQ 174
Db 219 -----ESPSCG-WTSSAVRGEMNVPLRAMPDPAQVY 250
QY 175 NUNGNVQPDSTAETAIYANTNSGTITPDLNPGVSMITNLVFDL--PDMTYGHVGQHSIV 232
Db 251 NPLGSIPTDSTSGHL--QMLKAMDVPIETALLMKNLSMNLRSPDF-----EQHRRLL 302
QY 233 ASMGFFGSD 241
Db 303 SSSYSGAD 311

RESULT 7

US-11-087-099-2610
; Sequence 2610, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2610
; LENGTH: 993
; TYPE: PRT
; ORGANISM: Pseudomonas syringae pv. tomato str. DC3000
US-11-087-099-2610

Query Match 7.0%; Score 90.5; DB 7; Length 993;
Best Local Similarity 23.6%; Pred. No. 3.8;
Matches 48; Conservative 33; Mismatches 65; Indels 57; Gaps 9;

QY 27 ILRFRGSLWA-----LIVFANTVYMRGNTPRNASTV---SATTSLVNSTNSQVAKQ 77
Db 734 ITHAGRGRLMADSLDSGVSPFTLPMHQGDITGISTFPGKQGTTRRLAQSISSKLRQ 793
QY 78 QNNSSTSPA-----HKSTNSLQAHQAATSSQSKLRITPHTYGVGDLEIRVNSL 130
Db 794 LEDRLTRAGLNLNHRVNTLATVQALASLTVNSTSL-----DSFRK--SPDARLFWL 846
QY 131 QOVKSVGYDGGTANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNVQPDSTAETAIY 190
Db 847 SQAR-----AEWM-----ST--ELADILAQLDVNGGQHRITFTGD-- 880
QY 191 ANTNSGTITPDLNPGVSMITNLV 213
Db 881 -----FVRLRPRISLTLSMV 895

RESULT 8

US-11-130-821-4
; Sequence 4, Application US/11130821
; Publication No. US20060019275A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: Cowan, David
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding Proteins Involved in Sensory
; FILE REFERENCE: 02307E-084210US
; CURRENT APPLICATION NUMBER: US/11/130,821
; CURRENT FILING DATE: 2005-05-16
; PRIOR APPLICATION NUMBER: US/09/361,630
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/094,464

/ PRIOR FILING DATE: 1998-07-28
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 4
/ LENGTH: 729
/ TYPE: PRT
/ ORGANISM: Mus sp.
/ FEATURE:
/ OTHER INFORMATION: mouse taste cell polypeptide (TCP) #2 amino acid
/ OTHER INFORMATION: sequence
US-11-130-821-4

Query Match 6.8%; Score 89.5; DB 7; Length 729;
Best Local Similarity 28.5%; Pred. No. 3.1;
Matches 37; Conservative 14; Mismatches 44; Indels 35; Gaps 6;

QY 42 NTVMIRGNTPRNAs-----TTSATSLVNSTN-SSQVAKQEQNS-----TSPAKSTNS 91
DB 243 NTVALDTGSPVSGADSDSPQTPTSTDSFKTSNLPQIALOPSHQMLPTSPHMPPLTS 302
QY 92 LQH-----AQHQAAATSSS-----QSLARYPHTYQK--VGDLEIR 126
DB 303 LQHSPPSTHASSGFTSSVHADPTLASTLPHPCQDMSLDLSPSTGSRHTHTSVPR 362
QY 127 VNSLQOVXSV 136
DB 363 INSNRFTKAV 372

RESULT 9
US-11-212-443-111
/ Sequence 111, Application US/11212443
/ Publication No. US20050287165A1
/ GENERAL INFORMATION:
/ APPLICANT: Scalco, Enzo
/ APPLICANT: Massignani, Vega
/ APPLICANT: Rappucci, Rino
/ APPLICANT: Pizza, Mariagrazia
/ APPLICANT: Grandi, Guido
/ TITLE OF INVENTION: Meningococcal Antigens
/ FILE REFERENCE: CHIR0159
/ CURRENT APPLICATION NUMBER: US/11/212,443
/ CURRENT FILING DATE: 2005-08-24
/ PRIOR APPLICATION NUMBER: US/09/302,626
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: PCT/IB99/00103
/ PRIOR FILING DATE: 1999-01-14
/ NUMBER OF SEQ ID NOS: 195
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 111
/ LENGTH: 600
/ TYPE: PRT
/ ORGANISM: Neisseria meningitidis
US-11-212-443-111

Query Match 6.8%; Score 87.5; DB 7; Length 600;
Best Local Similarity 20.4%; Pred. No. 3.6;
Matches 46; Conservative 42; Mismatches 120; Indels 17; Gaps 8;

QY 24 MNRIRFAGSLMALIVFANTYV--MIRGNTPRNASTVSAT--TSLVNSTNSQVAKQEQN 79
DB 1 MNRIRFAGSLMALIVFANTYV--MIRGNTPRNASTVSAT--TSLVNSTNSQVAKQEQN 79
QY 80 SSTSFAHSTNSLQAOHQAAAT-----TSSQSKLRYIPHTYQVGDLEIRVNSLQOVK 134
DB 58 EDLEEVVTRAPVLSFHSDEKGEKEVGAASNLTVYPDKNVLAAGITTLAAGNLKIK 117
QY 135 SVGYDGIETANGAFVINITIRNDSTPMEVVDGIFHLQNLNGNVYQPDSTAEIYANTN 194
DB 118 QNTNMENTMENTNASSFTYSLKDLGLIVETBKSFQANGKVAIIS-DTGLIAFPAKET 176
QY 195 SGT--IPT-DLNP-GVSMITTNLVFDMPDFTYGVHGOHYSIYASV 235

DB 177 AGTNGDPTVHLNGIGSTLTDTLAGSSASHVDAGNOSTHYTRAASI 221

RESULT 10
US-11-060-920-5
/ Sequence 5, Application US/11060920
/ Publication No. US20050244378A1
/ GENERAL INFORMATION:
/ APPLICANT: Kaufman, Paul L
/ APPLICANT: Liu, Xuyang
/ TITLE OF INVENTION: Method for Treating Glaucoma
/ FILE REFERENCE: 960296.00149
/ CURRENT APPLICATION NUMBER: US/11/060,920
/ CURRENT FILING DATE: 2005-02-18
/ PRIOR APPLICATION NUMBER: US 60/545,723
/ PRIOR FILING DATE: 2004-02-18
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: Patent In version 3.2
/ SEQ ID NO 5
/ LENGTH: 721
/ TYPE: PRT
/ ORGANISM: Clostridium botulinum
US-11-060-920-5

Query Match 6.8%; Score 87.5; DB 7; Length 721;
Best Local Similarity 21.5%; Pred. No. 4.6;
Matches 45; Conservative 39; Mismatches 96; Indels 29; Gaps 8;

QY 28 LRFAGSLMALIVFANTYVIRGNTPRNASTVSATTS-----LVNSTNSQVAK--QEQ 78
DB 312 LQLAGGIRP--VFSMSASANYSHWTONTSTVDDTGSSFSQGLINTGSAVINENIRY 369
QY 79 NSTSPARK--STNSLQAOHQAAATSSSQSL-RYI-PFHTYQVGDLEIRVNSLQOVK 134
DB 370 NTGAPVNTVPTTITVIDKQSVATIKQSLIGDYLMPGGTYPPIGSPPALNTMDQPS 429
QY 135 S-----VGYDGIETANGAFVINITIRNDSTPMEVVDGIFHLQNLNGNVYQPDSTAEIY 190
DB 430 SRLPINTNQLKSIDNGTWTLS-----QFGNFRAKNSNGNLVTDGNNMGPY 479
QY 191 ANTNSTGTPDLPNGVSMITTNLVFDMPDPF 219
DB 480 LGITKSTYASLTLSFSQGTQVAVAVAPNF 508

RESULT 11
US-11-052-554A-95
/ Sequence 95, Application US/11052554A
/ Publication No. US2005028866A1
/ GENERAL INFORMATION:
/ APPLICANT: Sachdeva, et al.
/ TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
/ FILE REFERENCE: 30853/40359A
/ CURRENT APPLICATION NUMBER: US/11/052,554A
/ CURRENT FILING DATE: 2005-02-07
/ PRIOR APPLICATION NUMBER: US 60/589,227
/ PRIOR FILING DATE: 2004-07-20
/ PRIOR APPLICATION NUMBER: IN 173/DEL/2004
/ PRIOR FILING DATE: 2004-02-06
/ NUMBER OF SEQ ID NOS: 763
/ SOFTWARE: Patent In version 3.3
/ SEQ ID NO 95
/ LENGTH: 1237
/ TYPE: PRT
/ ORGANISM: Helicobacter pylori J99
US-11-052-554A-95

Query Match 6.7%; Score 87; DB 7; Length 1237;
Best Local Similarity 21.1%; Pred. No. 11;
Matches 52; Conservative 37; Mismatches 70; Indels 88; Gaps 12;

QY 38 IVFANTVMIRGNTP-----RNASTVSATSLVNSTNSQVAKQEQ-----QNSSTSPA 85

GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchenev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spyrek, Kimberly
APPLICANT: Rattelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zernusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezei, Peter
APPLICANT: Patutajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Elesen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shinkets, Richard
APPLICANT: Guesv, Vladimir
APPLICANT: Vermet, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
PRIOR FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 141
LENGTH: 1940
TYPE: PRT
ORGANISM: Mus musculus
US-10-055-877-141

Query Match 6.6%; Score 85; DB 6; Length 1940;
Best Local Similarity 24.6%; Pred. No. 31;
Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;
QY 59 SATSLVNSTNSQVAKQEQNSSTSPAHKSTNSLQHAQQAATTSSQSKLRYIPHTY 118
DB 670 NANVLSKSGITPLHLAAGEPRVAVAEVLVQGAHV-----DAQTKMGTPLAVGC 721
QY 119 RVGDLIRVNSLQGVKSV-----GYDQIGETA-NGAFWVINITIRNDGSTPEVVDGI 170

DB 722 HYGNIKI-VNFILOHSAKNVAKTKNGYTAHLQAAQGHTHIINVLQNNAS-PNEL----- 775
QY 171 FHLQNLN-----YQPD-----TAEIYANTNSGTFPDNPGVSMTNLYPD 215
DB 776 ----TVNGNTALAIARRLGYISVDTLKVTEIMTTT---TTEKHQNNPETNEVLD 828
QY 216 MPD 218
DB 829 MSD 831
RESULT 15
US-11-096-568A-29105
Sequence 29105, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 29105
LENGTH: 946
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(946)
OTHER INFORMATION: Ceres Seq. ID no. 4263811
US-11-096-568A-29105
Query Match 6.6%; Score 84.5; DB 7; Length 946;
Best Local Similarity 21.7%; Pred. No. 13;
Matches 43; Conservative 32; Mismatches 82; Indels 41; Gaps 7;
QY 52 PRNASTVSATSLVNSTNSQVAKQEQNSSTSPAHKSTNSLQHAQQAATTSSQSKLRY 111
DB 59 PDNSSVLSLSNFDLSNSSFPLVLCNLTLESIDVSNRSLSPGCVTCERLIALKH 118
QY 112 IPFHT-----YQVLDLIRVNSLQ-QVSVGYDQIGETANGAFWVINITI-RN 158
DB 119 LNFSTNKRSTSPGRGFSKLAVLDPFHVLSGNVGDYGFGLVQLRS-----LNLSPNRL 173
QY 159 DGSPTMEVVDGIFHLQNLNGVNYQPDSTAEIYANTNSGTFITDLNPGVSMTNLYFMPD 218
DB 174 TGSVPVHLTKSLERL-----EVSDNSLSTIGIPEGIKDYQELT---LIDLSD 216
QY 219 FMTYGHVQHYSLVSMG 236
DB 217 NQLNG-----SIPSSLG 228

Search completed: April 7, 2006, 16:40:50
Job time : 18.448 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:08:38 ; Search time 247.035 Seconds

(without alignments)
368.172 Million cell updates/sec

Title: US-10-784-592-43_COPY_42_248

Perfect score: 1074
Sequence: 1 NTVMIRKGNTRRNASTVSAT.....YSLVSMGPRFGSDERTYALP 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21:*

- 1: geneseqp19808:*
- 2: geneseqp19908:*
- 3: geneseqp20008:*
- 4: geneseqp20018:*
- 5: geneseqp20028:*
- 6: geneseqp20038:*
- 7: geneseqp20048:*
- 8: geneseqp20058:*
- 9: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1074	100.0	207	9	ABE45583 A1cyclob
2	1074	100.0	248	9	ABE45569 A1cyclob
3	1074	100.0	248	9	ABE48755 A1cyclob
4	135	12.6	331	4	AAU37553 Staphyloc
5	135	12.6	331	4	AAU34416 Staphyloc
6	135	12.6	331	4	AAU37279 Staphyloc
7	135	12.6	331	6	ABU16122 Protein e
8	135	12.6	350	6	ABU11412 Staphyloc
9	113.5	10.6	6713	6	ABU15887 Protein e
10	107.5	10.0	2086	4	AAU34143 Staphyloc
11	107.5	10.0	5795	4	AAU37017 Staphyloc
12	105.5	9.8	922	8	ADJ34822 Xylanase
13	104	9.7	417	7	ABO70979 Pseudomon
14	103	9.6	2344	4	AAU37120 Staphyloc
15	101.5	9.5	337	5	AAU19444 Staphyloc
16	99	9.2	485	6	ABP56884 Staphyloc
17	99	9.2	2261	6	ABU18914 Pathogen
18	99	9.2	2271	6	ABU16000 Protein e
19	99	9.2	2271	6	ABU17274 Staphyloc
20	98.5	9.2	2283	6	ABP56876 Staphyloc
21	98.5	9.2	2398	6	ABU42252 Protein e
22	98.5	9.2	6281	4	AAU37403 Staphyloc
23	98.5	9.2	9535	6	ABU73008 Staphyloc
24	98.5	9.2	10498	6	ABU19119 Pathogen

25	97	9.0	194	6	ABU29785
26	95	8.8	982	4	ABE59353
27	95	8.8	982	4	ABE67401
28	94.5	8.8	498	1	AAU70302
29	94.5	8.8	632	7	ADU60974
30	94	8.8	349	3	ABE26877
31	94	8.8	1375	5	AAU98028
32	94	8.8	1375	5	AAU79288
33	94	8.8	1375	7	ADU93655
34	94	8.8	1375	9	ADU37278
35	94	8.8	1484	8	ADU34858
36	94	8.8	2712	6	ABU39146
37	93.5	8.7	498	4	AAU49641
38	93.5	8.7	603	8	ADU73414
39	93.5	8.7	627	7	ADU60976
40	93.5	8.7	633	2	AAU75774
41	93.5	8.7	661	5	ABG93298
42	93.5	8.7	954	4	ABE71703
43	93.5	8.7	1309	4	ABG30366
44	92.5	8.6	1228	6	ABU03505
45	92.5	8.6	1228	8	ADU39914

ALIGNMENTS

RESULT 1	ABE45583	standard, protein; 207 AA.
ID	ABE45583	
XX	ABE45583	
AC	ABE45583	
DT	22-SEP-2005	(first entry)
XX		
DE	Allicyclobacillus sp. mature functional polypeptide (amino acids 42-248).	
XX		
KW	Feedstuff; food; detergent; surfactant; pulp; functional polypeptide.	
XX		
OS	Allicyclobacillus sp.; DSM 15716.	
XX		
PN	US2005147983-A1.	
XX		
PD	07-JUL-2005.	
XX		
PF	23-FEB-2004; 2004US-00784592.	
XX		
PR	06-JAN-2004; 2004DK-00000010.	
XX		
PR	04-FEB-2004; 2004DK-00000165.	
XX		
PA	(NOVO) NOVOZYMES AG.	
XX		
PI	Willing R, Lassen SF, Ostergaard PR,	
XX		
DR	WPI. 2005-511773/52.	
XX		
PT	N-PSDB; ABE45582.	
XX		
PT	New functional polypeptides having function and amino acid sequence similar to known specific bacterial enzymes useful in industrial, research and household applications e.g. detergents and food.	
XX		
PS	Claim 36; Page; 83pp; English.	
XX		
CC	The present invention relates to functional polypeptides encoded by	ABU29785 Protein e
CC	polymucleotides comprised in the genome of Allicyclobacillus sp. DSM	ABE59353 Drosophil
CC	15716. The invention is useful in industrial, research and household	ABE67401 Drosophil
CC	processes such as in composition to prepare feed and food, in detergent	AAU70302 Sequence
CC	formulations and for treating lignocellulosic fabric and pulp. The	ADU60974 B. thurin
CC	present sequence is Allicyclobacillus sp. DSM 15716 mature functional	ABE26877 Glucosyl
CC	polypeptide. Note: This sequence is not shown in the specification	AAU98028 S. mutans
CC	constructed based on the amino acid positions provided in claim 36 of the	AAU79288 Streptoco
CC	specification.	ADU93655 Streptoco
XX		ADU37278 Streptoco
XX		ADJ34858 Xylanase
XX		ABU39146 Protein e
XX		AAU49641 Escherich
XX		ADU73414 Plant ful
XX		ADU60976 B. thurin
XX		AAU75774 Amino aci
XX		ABG93298 C. albica
XX		ABU1703 Drosophil
XX		ABG30366 Novel hum
XX		ABU03505 Angiogene
XX		ADU39914 Human myo
SQ	Sequence 207 AA;	

Query Match	100.0%;	Score 1074;	DB 9;	Length 207;
Best Local Similarity	100.0%;	Pred. No. 1e-89;		
Matches 207; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY 1 NTYVMIRGNTPRNASTVTSATTSVLVNSTNSQVAKOEONSSSTSPAKSTNSLQHQHQAAAT 60
Db 1 NTYVMIRGNTPRNASTVTSATTSVLVNSTNSQVAKOEONSSSTSPAKSTNSLQHQHQAAAT 60

QY TSSSSKLRIPFHHYGVKVDLEIKVNSLQVKSGYDGIETANGAPWVINITIRNDGS 120
61 TSSSSKLRIPFHHYGVKVDLEIKVNSLQVKSGYDGIETANGAPWVINITIRNDGS 120
Db

QY 121 TPMEVVDGIFHLQNLNGVYQPDSIAEIVANTNSGTI PTDLNPGVSMITNLVFDMDPDMT 180

Db 121 TPMEVVDGIFHLQNLNGVYQPDSIAEIVANTNSGTI PTDLNPGVSMITNLVFDMDPDMT 180

QY	181	181	207
YGHVGQHYSLVASMGFPGSDETTYALP	YGHVGQHYSLVASMGFPGSDETTYALP	YGHVGQHYSLVASMGFPGSDETTYALP	YGHVGQHYSLVASMGFPGSDETTYALP
Db	181	207	

RESULT 2
AEB45569
ID AEB45569 standard; protein; 248 AA.

AC	AEB45569;	
XX		
DT	22-SEP-2005	(first entry)
YY		

DE	Allyclobacillus sp. functional polypeptide, SEQ ID NO: 43.
XX	
KW	Feedstuff; food; detergent; surfactant; pulp; functional polypeptide.
VV	

OS	Key	Location/Qualifi
XX	Alicyclobacillus sp.; DSM 15716.	
FH		
ET		

FT	/label= signal_peptide
FT	1
FT	Misc-difference
FT	/note= "Encoded by GTG
FT	protein
	43 348

FT	/note="Alicyclobacillus sp. mature functional
FT	polypeptide"
XX	
DN	ttc2005147003-11

XX	07-JUL-2005.
PD	
XX	
PR	23-FEB-2004.

XX	06-JAN-2004; 2004DK-00000010.
PR	04-FEB-2004; 2004DK-00000165.
PR	
XY	

PA (NOVO) NOVOZYMES AS.
XX
PI Wilting R, Lassen SF, Ostergaard PR,
XX

DR	WPI; 2005-511773/52.
DR	N-PSDB; ABB45544.
XX	
PT	New functional polymer

PT similar to known specific bacterial enzymes useful in industrial,
 PT research and household applications e.g. detergents and food.
 XX
 PS Example 2. SEQ ID NO 43: 83nm; English

XX The present invention relates to functional
CC polynucleotides comprised in the genome
CC 15716 The invention is useful for

CC processes such as in composition to prepare feed and food, in detergent
CC formulations and for treating lignocellulosic fabric and pulp. The
CC present sequence is *Aticyclobacillus* sp. DSM 15716 functional

CC polypeptide.
XX
SQ Sequence 248 AA;

Query Match	100.0%	Score 1074	DB 9	Length 248
Best Local Similarity	100.0%	Pred. No. 1.3e-89		
Matches 207; Conservative	0	Mismatches	0	Indels 0
				Gaps 0

Qy 1 NTVMIRGNTPRNASTVSATSLVNSTNSQVAQEQNSSTSPARKSTNSLQHQHQAAAT 60

Db 42 NTVMIRGNTPRNASTVSATSLVNSTNSQVAQEQNSSTSPARKSTNSLQHQHQAAAT 101

QY 61 TSSSSKLTAYIPFHTYGVKVDLEIRVNSIQQVSVGVYDIEGTANGAFAVINITIRNDGS 120

102 TSSSGSKLTAYIPFHTYGVKVDLEIRVNSIQQVSVGVYDIEGTANGAFAVINITIRNDGS 161

Db

Qy 121 TPMEVVDGTFHLQNLNGNVYQPDSSTAIRIANTNSGTIPFDLPFGVSMITNLVFPMDPDMFT 180

Db 162 TPMEVVDGTFHLQNLNGNVYQPDSSTAIRIANTNSGTIPFDLPFGVSMITNLVFPMDPDMFT 221

QY	181	YGHVGQHYSLVASMGGFGSDETTYALP	207
Db	222	YGHVGQHYSLVASMGGFGSDETTYALP	248

RESULT 3
AEB48755
ID AEB48755 standard; protein; 248 AA
wt

AC	ABB48755;
XX	
DT	22-SEP-2005 (first entry)
XX	

DE	<i>Alicyclobacillus</i> sp. DSM 15716 functional polypeptide.
XX	
KW	Protein production; protein secretion; surfactant; feedstuff; food.

OS	Alicyclobacillus sp.; 'DSM 15716'.
XX	
FH	Key
em	Location/Qualifiers

FT	/label= signal_peptide
FT	Misc-difference 1
FT	/note= "Encoded by GTG"
FT	Protein 43 348

FT	/label= Mature_polypeptide
XX	
PN	WO2005066339-A2.
XX	

PD	21-JUL-2005.
XX	
PF	06-JAN-2005,
XX	

PR	06-JAN-2004;	2004DK-00000010.
PR	04-FEB-2004;	2004DK-00000165.
PR	23-FEB-2004;	2004US-00784592.
DD	05-FEB-2004	2004DY-00000000.

XX (NOVO) NOVOZYMES AS.
PA
XX
BT W41-4320 D 1 22000 CE

XX WPI; 2005-506869/51.
DR
DR N-PSDB; AEB48730.
YY

PT commercial research purposes, specifically for cleaning a cellulosic
PT fabric, preparing food or feed additive, and for treating lignolastic
PT materials and pulp.

PS Claim 3; SEQ ID NO 43; 151pp; English.
XX
CC The present invention is based on the finding of a strain of

CC Alicyclobacillus, namely Alicyclobacillus sp. DSM 15716, which grows at
CC low pH (approximately 4-5) and at high temperature (50-60 degrees C). It
CC is an object of the invention to identify and provide polypeptides CC
CC secreted from Alicyclobacillus sp. DSM 15716 because such polypeptides
CC may be used for industrial purposes and may also be produced in
CC industrially relevant processes and amounts. Thus, the invention provides
CC isolated mature functional polypeptides which are at least 90% identical
CC to, and exhibit and same function as a corresponding secreted polypeptide
CC ABB48738-ABB48762 from Alicyclobacillus sp. DSM 15716, and also provides
CC polynucleotides ABB48713-ABB48737 encoding such polypeptides. The
CC polypeptides are preferably enzymes having acid endoglucanase, acid
CC cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl
CC protease, serine protease, Hcra-like serine protease, disulfide
CC isomerase, gamma-D-glutamyl-L-amino acid endopeptidase, endo-beta-N-
CC acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase,
CC phytase, phospholipase C, polysaccharide deacetylase, xylanase
CC deacetylase or sulfite oxidase activity. A composition comprising such an
CC enzyme can be used in a detergent composition, or a food or feed
CC composition. Also claimed are: nucleic acid constructs, recombinant
CC expression vectors and host cells; methods of producing the polypeptides;
CC a storage medium suitable for use in an electronic device comprising
CC information of the amino acid sequence of a polypeptide of the invention
CC or the nucleotide sequence encoding the polypeptide; and a process
CC comprising employing the polypeptide, or polynucleotide encoding it, in
CC an industrial or household technical process. The present sequence is the
CC protein sequence of a functional polypeptide of the invention.

XX Sequence 248 AA;

XX Query Match % 100.0%; Score 1074; DB 9; Length 248;

XX Best Local Similarity 100.0%; Pred. No. 1.3e-89; Mismatches 207; Conservative 0; Indels 0; Gaps 0;

QY 1 NTVTMIRGTPRNASTVSAATTSIVNSTSSQVAKQONSSTSPARKSTNSLQHAQQAAT 60
DB 42 NTVTMIRGTPRNASTVSAATTSIVNSTSSQVAKQONSSTSPARKSTNSLQHAQQAAT 101
QY 61 TSSSOSKATYIPHTYKGVGDLIRVNSIQVKSVGVDIGETANCAFWVINTTINDGS 120
DB 102 TSSSOSKATYIPHTYKGVGDLIRVNSIQVKSVGVDIGETANCAFWVINTTINDGS 161
QY 121 TPEAEVVDGIFHLQNLNGVYQPDSTABIYANTNSGTFPTDLNKGVMNTNLVDPMDPMT 180
DB 162 TPEAEVVDGIFHLQNLNGVYQPDSTABIYANTNSGTFPTDLNKGVMNTNLVDPMDPMT 221
QY 181 YGHVGQHTSLVASMGEFFGSDETTYALP 207
DB 222 YGHVGQHTSLVASMGEFFGSDETTYALP 248

RESULT 4
AAU37553
ID AU37553 standard; protein; 331 AA.

AC AAU37553;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1723.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.

PT N-PSDB; AAS55412.

PS New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 13146; 511bp; English.

CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 331 AA;

XX Query Match % 12.6%; Score 135; DB 4; Length 331;

XX Best Local Similarity 31.0%; Pred. No. 0.0011; Mismatches 43; Indels 10; Gaps 4;

XX Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSIQVKSVGVDIGETANGAFWVINTTINDGSTPEAEVVDGIF 130
DB 81 HIGETVKNQDLEAVTVNSVETKSGVPSLAPFNAGIFVVAIVTKKNGKEALTTIDSSWF 140
QY 131 HLQNLNGVYQPDSTABIYAN-TNSGTFPTD-----LAPGVSMTNLVDPMDP 177
DB 141 KLKS-GDKTFEADNTGSMASQSDNGSIENSFPLORINDPSTAOGIKIVDVS 192

RESULT 5
AAU34416
ID AU34416 standard; protein; 331 AA.

AC AAU34416;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #692.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.

OS Staphylococcus aureus.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

[illegible]

```

Db      959 NAKTSANTTINGLGNLPLUOLQLOQDNLKHQVEQAQN--VGWNVCKDKGNTLNTNMAALRTSI 1016
Qy      116 RNDGSEPMEEVVDGIFFHLQNLGNVNYQPDSTAEIYVANTNSGTFPTDLNPGVSMNTNLVFDM 175
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1017 QNDWTT--KTSQNTYLDASDSKNNY---NTA---VNNANGVINATNNP--NMDANAINDM 1066

RESULT 10
AAU34143
ID      AAU34143 standard; protein; 2086 AA.
AC      AAU34143;
DT      14-FEB-2002 (first entry)
DE      Staphylococcus aureus cellular proliferation protein #419.
KW      Antisense; prokaryotic cellular proliferation protein; antibiotic;
       KW      antibacterial; drug design.
XX      Staphylococcus aureus.
OS      WO200170955-A2.
PN      27-SEP-2001.
PX      21-MAR-2001; 2001WO-US009180.
PR      21-MAR-2000; 2000US-0191078P.
PR      23-MAY-2000; 2000US-0206848P.
PR      26-MAY-2000; 2000US-0207727P.
PR      23-OCT-2000; 2000US-0242578P.
PR      27-NOV-2000; 2000US-0253625P.
PR      22-DEC-2000; 2000US-0257931P.
PR      16-FEB-2001; 2001US-0269308P.
XX      (ELIT-) ELITRA PHARM INC.
PA      Haeelbeck R, Ohlsen KU, Zyekind JW, Wall D, Trawick JD, Carr GJ;
PI      Yamamoto RT, Xu HH;
DX      MPI; 2001-611495/70.
DR      N-PSDB; AAS52002.
XX      New polynucleotides for the identification and development of
PT      antibiotics, comprise sequences of antisense nucleic acids.
XX      Example 3; SEQ ID NO 5639; 511pp; English.
XX      The invention relates to antisense inhibitors of genes essential to
CC      prokaryotic cellular proliferation, their use in identifying the genes,
CC      their use in the discovery of novel antibiotics, the essential genes
CC      themselves and the encoded proteins. The prokaryotes used are Baccharichia
CC      coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC      Pseudomonas aeruginosa, and Enterococcus faecalis. The invention is also
CC      useful for the identification of potential new targets for antibiotic
CC      development. The antisense nucleic acids can also be used to identify
CC      proteins used in proliferation, to express these proteins, and to obtain
CC      antibodies capable of binding to the expressed proteins. The proteins can
CC      be used to screen compounds in rational drug discovery programmes. The
CC      antisense nucleic acid sequence is also useful to screen for homologous
CC      nucleic acids which are required for cell proliferation in a wide variety
CC      of organisms. The present sequence represents an essential prokaryotic
CC      cellular proliferation protein. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX      SQ      Sequence 2086 AA;
       10.0%; Score 107.5; DB 4; Length 2086;
Query Match
Beet Local Similarity 24.2%; Pred. No. 4.3;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;
```

Oy	NTVVMIRRENTERNST--VSAATTSLVNSTSSQYAKQBNQNSTSPAHKSTNSLQHAQQA	58
Dd	1323 NQNYLDADSESRNATQAVTAABEGLNKQGTGANTSKADVDNALVTYRAKALNGAENLR	1382
Oy	59 ATTSSSQSKLVEIPFHTYGVKGDLEIRNNSLQQYKSV-GYDGIETANGAFWVINITIRN	117
Dd	1383 NTKTSATNTINGLPLRNLQLODNLKHQYBQAQNVAGVGDCKNTLANTMGLRTSLQN	1442
Oy	118 DGSITPMEVVDGIFHLQNLNGNVYQPDSTAEIYVANTSGTIPDLNPGVSMTTNLVFDM	175
Dd	1443 DNTT--KTSQNYLDASDSKNKNY---NTA---VANNANGVINVNTNP--NMDANANGM	1490
RESULT 11		
AAU37017	AAU37017 standard; protein, 5795 AA.	
XX	AAU37017;	
XX	14-FEB-2002 (first entry)	
XX	Staphylococcus aureus cellular proliferation protein #1187.	
XX	Antisense; prokaryotic cellular proliferation protein; antibiotic;	
KM	antibacterial; drug design.	
XX	Staphylococcus aureus.	
OS	MO200170955-A2	
XX	27-SEP-2001.	
PD	21-MAR-2001; 2001WO-US009180.	
PR	21-MAR-2000; 2000US-0191078P.	
XX	23-MAY-2000; 2000US-0206848P.	
PR	26-MAY-2000; 2000US-0207727P.	
PR	23-OCT-2000; 2000US-0242578P.	
PR	27-NOV-2000; 2000US-0253625P.	
PR	22-DEC-2000; 2000US-0257931P.	
PR	16-FEB-2001; 2001US-0269308P.	
XX	(BLIT-) BLITRA PHARM INC.	
PA	Haeelbeck R, Ohlsen KI, Zyekind JW, Wall D, Trawick JD, Carr GJ,	
PI	Yamamoto RT, Xu HH;	
PI	WPI; 2001-611495/70.	
XX	N-PSDB; AAS54876.	
DR	New polynucleotides for the identification and development of	
PT	antibiotics, comprise sequences of antisense nucleic acids.	
XX	Example 3; SEQ ID NO 12610; s11pp; English.	
XX	The invention relates to antisense inhibitors of genes essential to	
CC	prokaryotic cellular proliferation, their use in identifying the genes,	
CC	their use in the discovery of novel antibiotics, the essential genes	
CC	themselves and the encoded proteins. The prokaryotes used are Escherichia	
CC	coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,	
CC	Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also	
CC	useful for the identification of potential new targets for antibiotic	
CC	development. The antisense nucleic acids can also be used to identify	
CC	proteins used in proliferation, to express these proteins, and to obtain	
CC	antibodies capable of binding to the expressed proteins. The proteins can	
CC	be used to screen compounds in rational drug discovery programmes. The	
CC	antisense nucleic acid sequence is also useful to screen for homologous	
CC	nucleic acids. Which are required for cell proliferation in a wide variety	
CC	of organisms. The present sequence represents an essential prokaryotic	
CC	cellular proliferation protein. Note: The sequence data for this patent	
CC	did not form part of the printed specification, but was obtained in	
CC	electronic format directly from WIPO at	

CC	Seq ID	Accession	Score	DB	Length	Matches	Similarity	Pred. No.	Mismatches	Indels	Gaps
CC	XX	ftp.wipo.int/pub/published_pct_sequences									
XX	Sequence	5795 Aa;	10.04;	DB 4;	Length 5795;						
CC	Query Match		10.04;	Score 107.5;	DB 4;	Length 5795;					
CC	Best Local Similarity	24.28;	Pred. No. 17;								
CC	Matches	43;	Conservative 35;	Mismatches 87;	Indels 13;	Gaps 6;					
QY	1	NTVTMRGNTPRNAST--VSATTSLVNSTSSQVAKQBNSSPSPAHKSTNSLQHAQOA	58								
DB	3600	NQNTVLADDESKRVAAYTOAVTAAEGILNKQCGNTSKADVNALNTVTRAAALNGAENLR	3659								
QY	59	ATTSSSQSKARVYPRFHYTGKVGDLERVNSLQOQKSY-GRDGGERTNGAFWYINITRN	117								
DB	3660	NTKTSATNTTNGLEPNLTQLQKDNLKQVEQAQNVAGNVGKDKNTLNTMGALRTSIQ	3719								
QY	118	DGSTPMVEVNGIFHLQNLNGVYQDPSTAEIYANTNSGTIPDNLPGVMTNULVFPDM	175								
DB	3720	DNTT--KTSGNYLDASDSMKNNT---NTA---VNNANGVINTVNP--NMDAALINGM	3767								
RESULT 12											
ID	ADJ34822										
XX	ADJ34822	standard; protein; 922 AA.									
AC	ADJ34822;										
XX	ADJ34822;										
DT	22-APR-2004	(first entry)									
XX	22-APR-2004	(first entry)									
DE	xylanase from an environmental sample seq id 38.										
XX	xylanase from an environmental sample seq id 38.										
KW	antibacterial; fungicide; thermostable xyylanase activity;										
KW	dough conditioning; beverage production; nutritional supplement;										
KW	animal feed; lignin reduction; wood product; xylian; bacterial infection;										
KW	fungal infection; coccidiosis.										
OS	Unidentified.										
XX	Unidentified.										
XX	WO2003106654-A2.										
XX	WO2003106654-A2.										
PD	24-DEC-2003.										
XX	24-DEC-2003.										
XX	16-JUN-2003; 2003WO-US019153.										
FP	16-JUN-2003; 2003WO-US019153.										
XX	16-JUN-2003; 2002US-0389299P.										
PR	16-JUN-2003; 2002US-0389299P.										
XX	16-JUN-2003; 2002US-0389299P.										
PA	(DIVE-) DIVERSA CORP.										
XX	(DIVE-) DIVERSA CORP.										
PI	Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;										
PI	Beteghlaian A;										
DR	WPI; 2004-099016/10.										
XX	WPI; 2004-099016/10.										
DR	N-PSDB; ADJ34821.										
XX	N-PSDB; ADJ34821.										
PS	Claim 60; SEQ ID NO 38; 570pp; English.										
XX	Claim 60; SEQ ID NO 38; 570pp; English.										
CC	The invention describes an isolated or recombinant polypeptide (I),										
CC	having 50% or more identity to 190 300-1200 residue amino acid sequences										
CC	(S1), given in the specification, over a region of 100 or more residues										
CC	and the polypeptide as thermostable xyylanase activity. (I) is useful for:										
CC	dough conditioning; beverage production; as a nutritional supplement in										
CC	animal feed; reducing lignin in a wood or a wood product; and for										
CC	eliminating and protecting animals from a microorganism comprising xylian.										
CC	The polynucleotide (II) encoding (I) is useful for amplifying nucleic										
CC	acid encoding a polypeptide having a xyylanase activity which involves				</						

coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/publ/published_pct_sequences

SO Sequence 2344 AA;

Query Match 9.6%; Score 103; DB 4; Length 2344;
Best Local Similarity 25.5%; Pred. No. 13;
Matches 55; Conservative 18; Mismatches 71; Indels 72; Gaps 9;

QY 1 NTVTMRGNTPRASATYATSLVNST---NSQVAKQONSSTSPAHKS-----TNS 50
DB 101 NTGSETVGN--QNSTIIDASTADSTVTKNSSSV--QTSNSDVTSSKSENVSTSTNS 156

QY 51 LOHAHQAAATSSSQKRLRIPFHYTGKVGDDLEIRVNSLQOVKSVGYDGIGETANGAFW 110
DB 157 TSNQKRLTSTSTSK-----NTTSSDTSV----- 185

QY 111 INITIRNDSTPEWVDGIFHLQNLNGVYQPDSTAIFYANTNSGTLPTDLNPGVMTTN 170
DB 186 -----TSTSTBEPI-----NISTNOSTNS--NNTSQSTTPISANLAKTSTTS 226

QY 171 LVFDMDFMTYGHVGHQHSIVASMGFPQSDETTVAL 206
DB 227 TSTAPVKLRTFSRL-----AMSTFASNAATTAL 254

RESULT 15

AAE19444 ID AAE19444 standard; protein; 337 AA.

XX AAE19444;

XX AC 29-AUG-2003 (revised)
XX DT 31-MAY-2002 (first entry)

XX DE Hansenula polymorpha GPI cell wall anchor protein, Hppir2p.

XX KW Cell wall anchor protein, HpsED1, Hppir2, HpgAS1; Hptir1, HpcwP1, GPI;
XX glycosyl phosphatidyl inositol; surface expression system; industry;
XX medical; food; chemical; biochemical.XX OS *Pichia angusta*.

XX FH Key Location/Qualifiers

XX FT Peptide 1..18

XX FT Protein /label= Signal_peptide

XX FT 19..337

XX FT /note= "Mature Hppir2p"

XX PN MO200212509-A1.

XX PD 14-FEB-2002.

XX PF 27-JUL-2000; 2000WC-KR000819.

XX PR 26-JUL-2000; 2000KR-00042939.

XX PA (KOR-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

XX PI Choi E, Sohn J, Kim S;

XX XX

DR WPI: 2002-227157/28.
DR N-PSDB; AAD30793.

PT Novel cell wall protein genes HpsED1, Hppir2, HpgAS1, Hptir1, HpcwP1
PT derived from yeast *Hansenula polymorpha*, which encode the cell wall
PT proteins useful for construction of surface expression systems.

PS Claim 5; Page 112-114, 121pp; English.

XX The present invention relates to a cell wall protein gene derived from
CC *Hansenula polymorpha*, such as HpsED1, Hppir2, HpgAS1, Hptir1, HpcwP1. The
CC surface expression systems have numerous applications including
CC immobilisation of biocatalyst and large scale production of protein such
CC as enzymes, antigens, antibodies, etc. The surface expression systems
CC comprising Hptir1p, HpsED1p, HpgAS1p, HpcwP1p, are useful for exporting
CC glucose oxidase to cell surface. The surface expression systems have
CC biochemical industry. The surface protein derived from *H. polymorpha* are
CC highly effective in construction of surface expression systems and
CC development of biocatalyst application systems. The present sequence is
CC *Hansenula polymorpha* GPI (glycosyl phosphatidyl inositol) cell wall
CC anchor protein, Hppir2p. (Updated on 29-AUG-2003 to standardise OS field)

SO Sequence 337 AA;

Query Match 9.5%; Score 101.5; DB 5; Length 337;
Best Local Similarity 24.7%; Pred. No. 1.3;
Matches 54; Conservative 32; Mismatches 94; Indels 39; Gaps 8;

QY 14 ASTVSATSLVNSTNSQV-----KQONSSTSPAHKSTNSLOHAQHA--AT 60
DB 56 SSSVSASLVEKRVQVSDQIQATNTBKTSTKSTSTAAVAVSQITDQIQATAT 115

QY 61 TSSSQKRLRIPFHYTGKVGDDLEIRVNSLQOVKSVGYDGIGETANGAFWVINITIRNDG- 119
DB 116 TSSSSSSSKKTAAVVTOIGDQIQATSTSEKSTADAVVTOIGDQ---IQATSKSSST 172

QY 120 STEMEVY---DGIFH-----LQNLNGVYQPDSTAIFYANT-NSGTLPT 159
DB 173 STRADVSQLTGDQIQATSTSTRASASATTSQVISOISDQIQASSTASSTSSSTANG 232

QY 160 DLNPGVSMITT--LVFDMDFMTYGHVGHQHSIVASMGF 196
DB 233 DVTTSVCKKKGALANTLKGILYDSBGRIGSIVANROF 271

Search completed: April 7, 2006, 16:16:12
Job time : 248.535 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:18:42 ; Search time 22.7473 Seconds
(without alignments)
875.573 Million cell updates/sec

Title: US-10-784-592-43_COPY_42_248

Perfect score: 1074

Sequence: 1 NTVMIRGNTPRNASTVSAT.....YSLVSMGPRFGSDETTYALP 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	173.5	16.2	215	2	H84032	hypothetical prote
2	135	12.6	331	2	H89965	conserved hypothet
3	120.5	11.2	542	2	S64030	probable membrane
4	116	10.8	215	2	S77663	multicopy phage re
5	113.5	10.6	6713	2	B89921	hypothetical prote
6	105	9.8	389	2	S53975	probable membrane
7	100.5	9.4	202	2	B97130	uncharacterized se
8	100	9.3	394	2	A37853	flagellin, 40k - p
9	99	9.2	2271	2	P90073	hypothetical prote
10	97	9.0	846	2	AD2672	conserved hypothet
11	97	9.0	848	2	B97454	hypothetical prote
12	94	8.8	1375	2	UT0345	dextranase (EC
13	93.5	8.7	498	1	FLBC	flagellin - Zscher
14	93.5	8.7	784	2	S26638	SPR-1 protein - hu
15	93	8.7	982	2	T13653	hypothetical prote
16	92.5	8.6	535	2	B84443	hypothetical prote
17	92.5	8.6	1228	2	A57384	multimerin, endoch
18	91.5	8.5	1449	2	T30552	glucosyltransferas
19	91	8.5	452	2	T40634	hypothetical prote
20	90.5	8.4	233	2	H86925	probable lipoprote
21	90.5	8.4	1365	2	T30822	limp protein - Myc
22	90	8.4	687	1	B69649	beta-galactosidase
23	90	8.4	1311	2	C84528	hypothetical prote
24	89.5	8.3	820	2	T17519	cell surface anti
25	89.5	8.3	1449	2	T30857	glucosyltransferas
26	89	8.3	443	2	T40167	hypothetical prote
27	89	8.3	453	2	T48240	hypothetical prote
28	88.5	8.2	365	2	D70043	hypothetical prote
29	88.5	8.2	511	2	C86798	prophage p13 prote

30	88.5	8.2	633	2	C32053	paraportal crystal
31	88.5	8.2	1770	2	S56221	hypothetical prote
32	88	8.2	1051	2	T18351	limp protein - Myc
33	87.5	8.1	385	2	T38113	hypothetical serin
34	87	8.1	345	2	B83737	pectate lyase BH06
35	87	8.1	817	2	T22442	hypothetical prote
36	86.5	8.1	410	2	B35670	protein-tyrosine k
37	86.5	8.1	493	2	G90604	hypothetical prote
38	86.5	8.1	544	2	B97166	flagellar basal bo
39	86.5	8.1	666	2	S50452	hypothetical serin
40	86	8.0	349	2	T41394	hypothetical prote
41	86	8.0	518	2	T09354	hypothetical prote
42	86	8.0	1121	2	S54504	hypothetical prote
43	85.5	8.0	532	2	B35621	spore germination
44	85.5	8.0	1237	2	D71850	probable outer mem
45	85.5	8.0	1277	2	S70306	hypothetical prote

ALIGNMENTS

```

RESULT 1
H84032
hypothetical protein BH3064 [imported] - Bacillus halodurans (strain C-125)
C:/Date: 01-Dec-2000 #sequence revision 01-Dec-2000 #text_change 09-Jul-2004
C:/Accession: H84032
R:/Takami, H.; Nakarone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiki
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H84032
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-215 <STO>
A/Cross-references: UNIPROT:Q9K8R2; UNIPARC:UPI000000C40B3; GB:AP001517; GB:BA000004; NII
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH3064

Query Match          16.2% Score 173.5; DB 2; Length 215;
Best Local Similarity 30.2% Pred. No. 1.6e-06;
Matches 55; Conservative 32; Mismatches 80; Indels 15; Gaps 8;

QY      28 NSSQVAKQ-EQNSSTSPARKSTNSLQHQQAAT--TSSSQSKRTYIPHTYG--KVGSD 81
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      35 DSSSTAQEPEPAEBSVADQSENSEBPEBAGTEDETEEBAEEDPIAGIGELAKVGD 94
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      82 LEIRNSLQQVKSVCYDGIGETANGAFWVYINITIRNDGSTPEWVVDGIFHLQNLGNV-Y 140
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      95 VFTFANGSTAGSVG-DVLTAEAKGTFLVDVITIRKSGDSITVDSFPKLR--VQDVEY 151
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      141 QPDSATRIYNTNSGTFIPDLNPGVSMNTNLTVEFDM-PDMTGVGHQHSYLVSMGPFSS 199
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      152 DSDSSAGLYANEGADPFLTKLNPGLPEKRVFDPQDVLDSDDI---LLNVQTSFFGT 207
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      200 DE 201
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      208 QQ 209

RESULT 2
H89965
conserved hypothetical protein SA1618 [imported] - Staphylococcus aureus (strain N315)
C:/Species: Staphylococcus aureus
C:/Date: 10-May-2001 #sequence revision 10-May-2001 #text_change 09-Jul-2004
C:/Accession: H89965
R:/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89756; MUID:21311952; PMID:11418146

```


A:Reference number: S53969
A:Accession: S53975
A:Molecule type: DNA
A:Residues: 1-389 <CON>
A:Cross-references: UNIPROT: Q04951; UNIPARC:UPI00001356B9; EMBL: Z49212; NID: g798940; PIR: S53969
C:Genetics:
A:Gene: SGD: SCW10; MIPS: YMR305C
A:Cross-references: SGD: S0004921
A:Map position: 11R
C:Superfamily: family 17 glucosidase
C:Keywords: transmembrane protein
C:Keywords: transmembrane #status predicted <TM>
A:6-22/Domain: transmembrane #status predicted <TM>

Query March	9.8%	Score 105;	DB 2;	Length 389;
Best local Similarity	22.9%	Fred. No. 0.82;		
Matches	53;	Conservative	40;	Mismatches 90;
			Indels 48;	Gaps 8;
QY	5	MIRGNT-----	PRASIVSATTSS-----	LVNSTSSQVAKOEONS-----
		::: ::	::: ::	::: ::
Db	43	SVGNSGRTIVPVNENAVATTSS	PAVASQATSTLTERTSANVTSSQ	QSTLQSSRA 102
QY	40	-----STPAKSTNSLOHQAQ	AATSSQSK-LRIIPFTYKXVDLE	IVNSLQ 91
		::: ::	::: ::	::: ::
Db	103	STVSSSTSSSSSSSTSS	SASSSIASGAKGITYSPYNDG	SCKSTPAQVADLEQ 162
QY	92	VKSVGYDQIGETANGAFV	INITIRNDGSTMEAEVDDG	FHLQNTNGNVYQPDSTARIYAN 151
		::: ::	::: ::	::: ::
Db	163	L--IGFNDIRIRIYGD	CSQSEVNVLOAKTSSQCLFL--	GIYYDKQDAVDITKSAVESTGS 218
QY	152	TNSGT---IPFDLNP	GVSMTTNLVPMDFPTYG	GHVGHYSLVASMGFGS 199
		::: ::	::: ::	::: ::
Db	219	WDDITTVSVGNELVNG	SGATTQVGE-----	YVSTAKALTSAGTYS 261

RESULT 7
E97130
uncharacterized secreted protein, homolog YXKC *Bacillus subtilis* [Imported] - Clostridium
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: E97130
R/Noiling, J.; Britton, G.; Omelchenko, M.V.; Matkova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4833-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: E97130
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-202 <KUR>
A/Cross-References: UNIPROT:Q97HY8, UNIPARC:UPI00000CA32C; GB:AE001437, PIDD:AAK79832.1
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
C/Gene: CAC1868

[illegible]

RESULT 8
A37853

flagellin, 40K - *Pseudomonas aeruginosa*
 C|Species: *Pseudomonas aeruginosa*
 C|Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 26-Aug-1999
 C|Accession: A37853
 R|Totten, P.A.; Lory, S.
 J. Bacteriol. 172, 7188-7199, 1990
 A|Title: Characterization of the type 4 flagellin gene from *Pseudomonas aeruginosa* PAK.
 A|Reference number: A37853; MUID:91072275; PMID:2123866
 A|Accession: A37853
 A|Status: Preliminary
 A|Molecule type: DNA
 A|Residues: 1-394 <TOT>
 A|Cross-references: UNIPARC:UP1000016FCA1; GB:M57501; NID:g151225; PIDN:AAA63458.1; PID
 C|Superfamily: flagellin

Query Match	9.3%;	Score 100;	DB 2;	Length 394;
Best Local Similarity	24.4%;	Pred. No. 2.1;		
Matches	51;	Conservative 32;	Mismatches 72;	Indels 54; Gaps 10;
QY	NTPNASTVSA-----	TTSLVNS-----	TTNSOVAKOENBST	41
DB	13 NTGNLNNSSSALNTSLQRLSTGSRINSADODAGLQIANLNTSQVNGLANVATNANNGI			72
QY	42 SPA-----HKSTNSIQAHQH--QAATTSSQSKLRYIPFTHYGKVGDLLEIRVNSIQ			91
DB	73 SLAQTASGALQOOSTNIIQRMEDLSIQSANGSNSNSERTALN-----GEAKOQLOKEIDR			125
QY	92 VKSYGVGIGETAGTAFVIVINITIRNDGSTPMKVVVD--GIFFL--QNLNGANYOPDSTABI			148
DB	126 ISNTTTCGKRLDSDGSRFVASFQV---GSANAEIISVIGIDEMASLSLNGTYFKADGGAV			182
QY	149 YANTNSGTTPTDLNPGVS--MTTNLVFDM			175
DB	183 TAAATASGTV--DIAIGITGSGAAVAVVKDM			209

```

RESULT 9
F90073
hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
C|Species: Staphylococcus aureus
C|Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C|Accession: F90073
R|Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu-
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A|Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A|Reference number: A89758; MUID:21311952; PMID:11418146
A|Accession: F90073
A|Status: preliminary
A|Molecule type: DNA
A|Residues: 1-2271 <R>
A|Cross-references: UNIPROT:Q990Y4, UNIPARC:UP1000000CABB3, GB:BA000018, PDB:q13702612;
A|Experimental source: Strain N315
C|Genetic8:
;Gene: SA2447

```

Query Match	9.2%	Score 99	DB 2	Length 2271
Beet Local Similarity	24.5%	Pred. No. 25		
Matches 53	Conservative 21	Mismatches 70	Indels 72	Gaps 9

QY	1	NTVYMGINPNRNASTVATSLVNSR	-----NSGQVAKQKONSETPAHS	-----TNS	50
DB	101	NTQSETVGN--QNSTTEASTSTADSTSVTKNSSV--QTSNBDTVSSEKRYTSTNS	156		
QY	51	LOHAQOAAATSSSSQSLRYIPFTYGVGLIRVNSLQOVKSVGVDGIGETANGAFWV	110		
DB	157	TSNQEKLTJTSBSTSK-----NTSSSDTQSV	186		
QY	111	INTIRNDGSTPEWVDIGIFHLQNLNGVYQOPDSTAETIVANTNSGTLPTDLNPGVMTTN	170		
DB	187	-----STSTSEQPI-----NTSTINOSTAS--NTTSQSTPPSSVNLNKTSTTS	226		

Db

Qy 109 WVNIT-----INDDSTMEWVDGFIHLQINNGVQGPDSIAEYANT--NSGT 157
 Db 184 -VNIIRKQNGKYTYKEDGTQKN-----YALNNGKTFEFDGTALSNNTLPKKGNI 236
 Qy 158 PTDLPNGVSTNTNLVPDMPDFTYGHVGOHY 188
 Db 237 TNNDDNTNSFAQYQOVS-TDVANFEHY-DHY 265

RESULT 13

flagellin - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 A:Variety: strain K-12
 C:Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text_change 09-Jul-2004
 A:Accession: A37249; J00018; I41270; H64955; A28187
 R:Kuwajima, G.; Aasek, U.I.; Fujiwara, T.; Node, K.; Kondo, E.
 J. Bacteriol. 168, 1479-1483, 1986
 A:Title: Nucleotide sequence of the hag gene encoding flagellin of Escherichia coli.
 A:Reference number: A37249; MUID:87057066; PMID:353685
 A:Accession: A37249
 A:Molecule type: DNA
 A:Residues: 1-498 <K0>
 A:Cross-references: UNIPROT:P04949; UNIPARC:UPI00000000AD; GB:M14358; NID:g146311; PIDN:
 A:Experimental source: Strain K-12
 R:Hanatsuma, T.; Sakai, A.; Tomimaga, A.; Enomoto, M.
 Mol. Gen. Genet. 216, 44-50, 1989
 A:Title: Isolation and characterization of Escherichia coli hag operator mutants whose h
 A:Accession: J00018; MUID:89281489; PMID:265972
 A:Molecule type: DNA
 A:Residues: 1-284; 'L', 286-498 <HAN>
 A:Cross-references: UNIPARC:UPI000016F1F3; GB:X17440; NID:g41649; PIDN:CAA35488.1; PID:g
 A:Experimental source: strain K-12
 R:Szekely, E.; Simon, M.
 J. Bacteriol. 155, 74-81, 1983
 A:Title: DNA sequence adjacent to flagellar genes and evolution of flagellar- phase vari
 A:Reference number: I41270; MUID:83238225; PMID:6305924
 A:Accession: I41270
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-20 <RES>
 A:Cross-references: UNIPARC:UPI000016F1F5; GB:J01607; NID:g146315; PIDN:AAA2491.1; PID:
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perma, N.T.; Burland, V.; Riley, M.; CC
 A: Rose, D.J.; Mau, B.; Shaq, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: H64955
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-498 <BIAT>
 A:Cross-references: UNIPARC:UPI00000000AD; GB:AE000285; GB:U00096; NID:g1788229; PIDN:AA
 A:Experimental source: strain K-12, substrain MG1655
 R:Kuwajima, G.
 J. Bacteriol. 170, 3305-3309, 1988
 A:Title: Construction of a minimum-size functional flagellin of Escherichia coli.
 A:Reference number: A28187; MUID:86257060; PMID:3290204
 A:Accession: A28187
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-15, 140--402, 489-498 <KU2>
 A:Cross-references: UNIPARC:UPI0000115532; GB:M21445; NID:g146313; PIDN:AAA23951.1; PID:
 C:Comment: This is the principal protein component of the flagellar filament, which is a
 C:Genetics:
 A:Gene: flhC; hag
 A:Map position: 42 min
 C:Superfamily: flagellin
 C:Keywords: flagellum

Matches	59; Conservative	33; Mismatches	99; Indels	81; Gaps	11;						
QY	12	RNASVTGATTSLVN	STN	SSQVAKQEQNS	SSSPAHKSTNSLOHQAHOAA	TSSSGSKRYI 71					
Db	53	RFTSNIKGLTQAA	RANNDGISVAQ	TREGALISEINN	LQKRVRELTVQA	TGTGNSBSDDSSI 112					
QY	72	P-----	FTATGKVDL	LRVNS-----	LOQV--	KSQVYDG 99					
Db	113	QDEIKSR	LDIEDRVS	GGQTGFEN	GNVNLAKNS	SMKIQVGANDNQIT	ITIDLKQIDATK	LDGDD 172			
QY	100	IG-----	ETANGAF	FWINTIT	-----	RNDGSI	PRMEVVDGIFHLQ	NLNG 137			
Db	173	PSVKN	NDVTTSAP	VPVAFGA	TTNNIKLTGIT	TLSTEATPTG	GTNPASIEG	VY--TDNG 229			
QY	138	NVYQP-----	DSTAEI	AYANT--	NSG	TIP-----	TDLPNG	VSMT-----	TNLVPD 174		
Db	230	NDVYAKITG	DDNGDK	KYAAV	TVVANDG	IVTMA	TGATPANA	TVTDANT	TKATIT	TSGGTP	QVID 289
QY	175	MPDEMTY	GHVGH	QHSLS	VASM	GFGS	DE	TYAL 206			
Db	290	NTAGS	ATANLGA--	VS	LVKLD	OSK	ND	DDIYAL 320			

RESULT 14

SPR-1 protein human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S26638
 R:Hagen, G.; Mueller, S.; Beato, M.; Ssnake, G.
 Nucleic Acids Res. 20, 5519-5525, 1992
 A:Title: Cloning by recognition site screening of two novel GT box binding proteins: a
 A:Reference number: S26638; MID:93087156; PMID:1454515
 A:Accession: S26638
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-784 <HAG>
 A:Cross-references: UNIPROT:Q02446; UNIPARC:UPI0000135D7E; EMBL:X66561; NID:g38419; PIDD
 C:Genetics:
 A:Gene: GDB:SP4, SPR-1
 A:Cross-references: GDB:136781
 A:Map position: 2q31.2q31
 A:Keywords: DNA binding; transcription regulation

[illegible]

RESULT 15

hypothetical protein 9557.2 - fruit fly (*Drosophila melanogaster*)
C|Species: *Drosophila melanogaster*
C|Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C|Accession: T113553
R|Ferriz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.
submitted to the EMBL Data Library, April 1999
A|Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.
A|Reference number: Z17694
A|Accession: T113553
A|Status: preliminary; translated from GB/EMBL/DDBL
A|Molecule type: DNA
A|Residues: 1-982 <PER>
A|Cross-references: UNIPARC:UPI0000128461; EMBL:AL021728; NID:e1355643; PID:e1301389; PT
A|Genetics:

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:09:09 ; Search time 135.574 Seconds
(without alignments)
1077.233 Million cell updates/sec

Title: US-10-784-592-43_COPY_42_248

Perfect score: 1074

Sequence: 1 NTVMIRGMTPRNASTVSAT.....YSLVSMGPFSGSDFTTALP 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trcembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	173.5	16.2	215	2	09K8E2_BACHD
2	139	12.9	331	2	06G8D3_STAAS
3	139	12.9	331	2	08NVZ1_STAW
4	135	12.6	331	2	06GFQ7_STAAC
5	135	12.6	331	2	05HEX9_STAAC
6	135	12.6	331	2	07A4Z0_STAW
7	135	12.6	331	2	09GT70_STAW
8	121.5	11.3	202	2	04MIR5_BACCE
9	120.5	11.2	542	1	SCW11_YEAST
10	119.5	11.1	170	2	08RB32_CHLTE
11	116	10.8	215	2	050442_MYCSC
12	113.5	10.6	6713	2	0931R6_STAW
13	113.5	10.6	6713	2	099U54_STAW
14	110	10.2	387	2	08G265_PSEAB
15	106.5	9.9	10746	2	06G6X3_STAR
16	105	9.8	389	1	SCW10_YEAST
17	105	9.8	2370	2	082RE3_STAW
18	104.5	9.7	766	2	054PB6_DICDI
19	104	9.7	329	2	06SYW1_PASMU
20	104	9.7	387	2	08G266_PSEAB
21	104	9.7	393	1	FLICA_PSEAB
22	104	9.7	394	2	053ZR9_PSEAB
23	104	9.7	394	2	053ZS1_PSEAB
24	103.5	9.6	923	2	07UKM5_RHOBA
25	103	9.6	713	2	06PES7_ACTIAD
26	102.5	9.5	1259	2	08G9X4_DICDI
27	102.5	9.5	9904	2	08NMQ6_STAW
28	100.5	9.4	202	2	097HY8_CIOAB
29	99.5	9.3	696	2	054Y58_DICDI
30	99	9.2	333	2	P72126_PSEAB
31	99	9.2	340	2	P72123_PSEAB

32	99	9.2	355	2	05Z159_NOCFA
33	99	9.2	1221	2	054LUB_DICDI
34	99	9.2	2261	2	05HC93_STAAC
35	99	9.2	2271	2	07A362_STAW
36	99	9.2	2271	2	099QY4_STAW
37	99	9.2	2275	2	06G620_STAAS
38	99	9.2	2275	2	08NUJ3_STAW
39	98.5	9.2	10498	2	05HPV8_STAAC
40	98	9.1	1128	2	07NZ14_CHRVO
41	98	9.1	1145	2	09UR03_CANAL
42	98	9.1	1298	2	08G647_DICDI
43	97.5	9.1	1359	2	08RG66_DICDI
44	97.5	9.1	1551	2	08IS13_DICDI
45	97.5	9.1	1557	2	054FP3_DICDI

ALIGNMENTS

RESULT 1					
ID	09K8E2_BACHD				
AC	09K8E2;				
DT	01-OCT-2000 (TREMblrel. 15, Created)				
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)				
DT	01-JUN-2003 (TREMblrel. 24, Last annotation update)				
DB	BH3064 Protein.				
GN	OrderedLocustNames=BH3064;				
OC	Bacillus halodurans.				
OS	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxId=86665;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C-125 / JCM 9153;				
RX	MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;				
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,				
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,				
RA	Horikoshi K.;				
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus				
RT	halodurans and genomic sequence comparison with Bacillus subtilis."				
RL	Nucleic Acids Res. 28:4317-4331(2000).				
DR	EMBL; BA000004; BAB06783.1; -; Genomic_DNA.				
DR	PIR; H84032; H84032.				
KW	Complete proteome.				
SQ	SEQUENCE 215 AA; 23072 MW; 6A05DASDE7DC6358 CRC64;				
Query Match					
Best local similarity 30.2%; Pred. No. 6.4e-06;					
Matches 55; Conservative 32; Mismatches 80; Indels 15; Gaps 8;					
QY	28 NSSQVAKQ-EQNSSTSPAHKSTNSLQHAQHOAA--TSSSQKLRYPHTYG---KVGCD 81				
DB	35 DSSRFAQPEPRBAEVSADQSENBSBPEBAEGTEDEYSBAEEDPIAGLGBALKVGD 94				
QY	82 LEIRNNSQQQKSVYDIDIGETANCAFWVITITINDSTPEWYVDGFLHNLNNGV-Y 140				
DB	95 VFTFANGSTGASVG-DVLTAAKAKTFLIVVTITNBSDSITVDSFFPKK--GVDEY 151				
QY	141 QPDSTAEIYATNAGSTIPTDLNPGVSMNTNLVFPDP--DFMTYGHGQHYSLVASMGPFS 199				
DB	152 DSDSAGXLYNAGDFFITKANFGLELPKRYVFDVFPQVLDSDDI-----LNVQGFPGT 207				
QY	200 DB 201				
DB	208 QQ 209				
RESULT 2					
ID	06G8D3_STAAS				
AC	06G8D3; STAAS PRELIMINARY;				
DT	05-JUL-2004 (TREMblrel. 27, Created)				
DT	05-JUL-2004 (TREMblrel. 27, Last sequence update)				


```
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN OrderedLocustNames=SAS1720;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Baeson N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571857; CAG43524.1; -; Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 331 AA; 35855 MW; 753AFB5FDAF50CF9 CRC64;

Query Match
Best Local Similarity 12.9%; Score 139; DB 2; Length 331;
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSLQOVKSVGYDIGETANGAFWVITIRNDGSTPMEVVDGIF 130
DB 81 HKIGETVKNGLDEVTVNSVETKSVGPSIAPTNAGTFVADVITIKNGKKAALTIDSSWF 140
QY 131 HLQNINGVYQPDSTAEIYANTN-SGTIPTD-----LNPVSMNTNLVFDMPD 177
DB 141 KTKS-GDKTFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIYFDVSR 192

RESULT 3
Q8NVZ1_STAAM
ID Q8NVZ1_STAAM PRELIMINARY; PRT; 331 AA.
AC Q8NVZ1;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein MW1738.
GN OrderedLocustNames=MW1738;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naiml T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramoto K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL: BA000033; BAB95603.1; -; Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 331 AA; 35855 MW; 753AFB5FDAF50CF9 CRC64;

Query Match
Best Local Similarity 12.9%; Score 139; DB 2; Length 331;
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSLQOVKSVGYDIGETANGAFWVITIRNDGSTPMEVVDGIF 130
DB 81 HKIGETVKNGLDEVTVNSVETKSVGPSIAPTNAGTFVADVITIKNGKKAALTIDSSWF 140
QY 131 HLQNINGVYQPDSTAEIYANTN-SGTIPTD-----LNPVSMNTNLVFDMPD 177
```

```
DB 141 KTKS-GDKTFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIYFDVSR 192

RESULT 4
Q6GF07_STAAR
ID Q6GF07_STAAR PRELIMINARY; PRT; 331 AA.
AC Q6GF07;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN OrderedLocustNames=SAR1880;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Baeson N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holtroyd S., Jagels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571856; CAG40870.1; -; Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 331 AA; 35850 MW; 82E9C61D2D59066 CRC64;

Query Match
Best Local Similarity 12.6%; Score 135; DB 2; Length 331;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSLQOVKSVGYDIGETANGAFWVITIRNDGSTPMEVVDGIF 130
DB 81 HKIGETVKNGLDEVTVNSVETKSVGPSIAPTNAGTFVADVITIKNGKKAALTIDSSWF 140
QY 131 HLQNINGVYQPDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 177
DB 141 KTKS-GDKTFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIYFDVSR 192

RESULT 5
Q5HEX9_STAAC
ID Q5HEX9_STAAC PRELIMINARY; PRT; 331 AA.
AC Q5HEX9;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Conserved domain protein, putative.
GN OrderedLocustNames=SACOL1847;
OS Staphylococcus aureus (strain COL).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=93062;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15774886; DOI=10.1126/JB.187.7.2426-2438.2005;
RA Gill S.R., Forts D.E., Archer G.L., Mongodin E.F., Deboy R.T.,
RA Ravel J., Paulsen I.T., Kolonay J.F., Brinkac L.M., Beanan M.J.,
RA Dodson R.J., Daugherty S.C., Madupu R., Angiuoli S.V., Durkin A.S.,
RA Haft D.H., Vamathevan J.J., Khouri H., Utterback T.R., Lee C.,
RA Dimitrov G., Jiang L., Qin H., Weidman J., Tran K., Kang K.H.,
RA Hance I.R., Nelson K.B., Fraser C.M.;
RT "Insights on evolution of virulence and resistance from the complete
RT genome analysis of an early methicillin-resistant Staphylococcus
RT aureus strain and a biofilm-producing methicillin-resistant
RT Staphylococcus epidermidis strain.";
RL J. Bacteriol. 187:2426-2438(2005).
```


DR	EMBL:	CP000046;	AAM36865.1;	-;	Genomic_DNA.
DR	TIGR:	SACOL1847;	-		
KW	Complete	proteome.			
SO	SEQUENCE	331 AA;	35864 MW;	F4347361D2D59472	CRC64;
<hr/>					
	Query Match		12.6%;	Score 135;	DB 2; Length 331;
	Best Local Similarity		31.0%;	Pred. No. 0.014;	
	Matches	35;	Conservative	25;	Mismatches 43; Indels 10; Gaps 4
Oy	74	HTYG---KVGDLIEIRVNSLQOVKS	VGDGIGETANGAFYVINITTRNDGSTPEWEVDGIF	130	
Dd	81	HKIGETVNKGDLLEVYNVSVEIKSVSPSLAP	TNAGKIIPVADVTIKNCKEALTTIDSSMF	140	
Oy	131	HLQNLGNVYQDPSTAEIYAN-TNSGTIPTD-----LN	GVSMTNLVDPMDP	177	
Dd	141	KLKS-GDKTFEADNTGMSMANSODNGSIENSF	LQRINDSTAQGIIVDSVSE	192	
<hr/>					
	RESULT 6				
	Q7A420 STAA	N PRELIMINARY;	PRT;	331 AA.	
ID	Q7A420;				
AC	Q7A420;				
DT	05-JUL-2004 (TrEMBLrel. 27,	Created)			
DT	05-JUL-2004 (TrEMBLrel. 27,	Last sequence update)			
DE	Hypothetical protein SAJ618.				
DN	OrderedLocustNames=SAJ618;				
OS	Staphylococcus aureus (strain N315).				
OC	Bacteria; Firmicutes; Bacillales;	Staphylococcus.			
OX	NCBI_TaxId=158879;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [Large scale genomic DNA].				
RX	MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;				
RA	Xukoda M., Ohra T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,				
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,				
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,				
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Katto C.,				
RA	Seikimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,				
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,				
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;				
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus				
RT	aureus."				
RL	Lancet 357:1225-1240(2001).				
RD	EMBL; BA000018; BAB42886.1; -;	Genomic DNA.			
KW	Complete proteome; Hypothetical protein.				
SO	SEQUENCE	331 AA;	35892 MW;	F2DDAA94971AIAAF	CRC64;
<hr/>					
	Query Match		12.6%;	Score 135;	DB 2; Length 331;
	Best Local Similarity		31.0%;	Pred. No. 0.014;	
	Matches	35;	Conservative	25;	Mismatches 43; Indels 10; Gaps 4
Oy	74	HTYG---KVGLIEIRVNSLQOVKS	VGDGIGETANGAFYVINITTRNDGSTPEWEVDGIF	130	
Dd	81	HKIGETVNKGDLLEVYNVSVEIKSVSPSLAP	TNAGKIIPVADVTIKNCKEALTTIDSSMF	140	
Oy	131	HLQNLGNVYQDPSTAEIYAN-TNSGTIPTD-----LN	GVSMTNLVDPMDP	177	
Dd	141	KLKS-GDKTFEADNTGMSMANSODNGSIENSF	LQRINDSTAQGIIVDSVSE	192	
<hr/>					
	RESULT 7				
	Q99T70 STAA	N PRELIMINARY;	PRT;	331 AA.	
ID	Q99T70 STAA	N PRELIMINARY;	PRT;	331 AA.	
AC	Q99T70;				
DT	01-JUN-2001 (TrEMBLrel. 17,	Created)			
DT	01-JUN-2001 (TrEMBLrel. 17,	Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24,	Last annotation update)			
DE	Hypothetical protein.				
DN	OrderedLocustNames=SAJ1800;				
OS	Staphylococcus aureus (strain Mu50 / ATCC 700699).				
OC	Bacteria; Firmicutes; Bacillales;	Staphylococcus.			
OX	NCBI_TaxId=158879;				
RN	[1]				

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,
RA Mizutani H., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hiraoka H., Kihara S., Goto S., Yabuzaki J.,
RA Kishimoto M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratazu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus".
RL Lancet 357:1225-1240(2001).
DR EMBL/BA000017; BAB57962.1; -; Genomic_DNA.
DR PIR/H89965; H89965.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 331 AA; 35892 MW; P25DAA94971AAAF CRC64;

Query Match 12.6%; Score 135; DB 2; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.014;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

Qy 74 HTVG---KVGDLERNVSLDQVSVGVDGIGETANGAFWYINTIRNDGSTPEVVDGIF 130
Db 81 HKIGETVKNGLDETVAVSVEIKSVGSLAPTNAGIFVADVIRKKGKBAALIDSMF 140
131 HLOWANGVQPDSTAIYAN-TNSGPIFPD-----INPGVMTWNVFPDMPD 177
141 KLAS-GDKTRBADVTGSMANQSDNGSIENSFLQRIINPSTAGKIVPDVSE 192

RESULT 8
Q4MIR5_BACCE
Q4MIR5_BACCE PRELIMINARY; PRT; 202 AA.
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Hypothetical protein.
DE ORFNames=BCE_G9241.pBCLin29_0020;
GN Bacillus cereus G9241.
OS Bacillus cereus G9241.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_Taxid=269801;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=G9241;
RC PubMed=1515910; DOI=10.1073/pnas.040241401;
RA Hofmeister A.R., Ravel J., Raske D.A., Chapman G.D., Chute M.D.,
RA Marston C.K., De B.R., Sacchi C.T., Fitzgerald C., Mayer L.W.,
RA Melden M.C.J., Priest F.G., Barker M., Jiang L., Cer R.Z.,
RA Ralston J., Peterson S.N., Weyant R.S., Galloway D.R., Read T.D.,
RA Popovic T., Frazer C.M.;
RT "Identification of anthrax toxin genes in a Bacillus cereus associated
RT with an illness resembling inhalation anthrax."
RL Proc. Natl. Acad. Sci. U.S.A. 101:8449-8454(2004).
RI -!- CAUTION: The sequence shown here is derived from an
RI EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
RI preliminary data.
DR EMBL/AE00100053; EAL12062.1; -; Genomic_DNA.
DR DR Hypothetical protein.
SQ SEQUENCE 202 AA; 22430 MW; 7AA489138004F80C CRC64;

Query Match 11.3%; Score 121.5; DB 2; Length 202;
Best Local Similarity 23.7%; Pred. No. 0.09;
Matches 37; Conservative 33; Mismatches 77; Indels 9; Gaps 3;

Qy 28 NSSQVAKQBNQSTSPAHKSTNSLOAHQAOAATSSQSGTLKRIPIHTYGVKGD---LRI 84
Db 25 NSTKDAEETKTAETKQKQAKQSTKQDEKKEAKSBEPKKE--PKKSLKGGESSKXVI 82
85 RVNSLDQVSVGVDGIGETANGAFWYINTIRNDGSTPEVVDGIFHLQNLNGNVQPD 144
83 AVGVSTVSTSVGGQVLSERKQGVFKVLEISTTNQDADITVDNSFKLVNDQDRBKYSST 142

```



```
Db 72 INQPDAAVLFDVITVRNDKREARTIAP--FKLIDENGAEYETSSNA--WSVDGSLGILD 127
QY 160 DLNPGVSMNTNVPFMDFTYGHVGOHSLVASMFPSSDE 201
Db 128 SLNPGYKRGYIVFDVFR-----GKHVXLEVGSGVSSDK 162

RESULT 11
050442 MYCSM
ID 050442_MYCSM PRELIMINARY; PRT; 215 AA.
AC 050442;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
DE Mpr.
GN Name=mpf;
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97000357; PubMed=843442;
RA Barom B.K., Hatfull G.F.;
RT "Characterization of Mycobacterium smegmatis gene that confers
RL Mol. Microbiol. 21:159-170(1996).
DR EMBL; U50355; AAB41651.1; -; GenomDc_DNA.
DR PIR; S77663; S77663.
SQ SEQUENCE 215 AA; 22960 MW; 7B4786D30BD6AB86 CRC64;

Query Match
Best Local Similarity 24.5%; Score 116; DB 2; Length 215;
Matches 37; Conservative 33; Mismatches 59; Indels 22; Gaps 5;

QY 36 EONSSTPAKSTNSLQHQHOA-----TTSSSOSKLRYPHTYGVGDLEIRN 87
Db 52 DDKQDTTAPL--NTATYQAAPALAPTKVEPTTINTIPV-----DKEEFVVR 99
QY 88 SLQD-VKSVGVYDG-IGETANGAFWVINTITRNDGSTPEVVDGIFHLQNLNGVYQPDST 145
Db 100 SVEPLSEVGNPFLNQKQKQGFVITLVSQNIQGRPGQSFPSNKLPTBRSRFTDTS 159
QY 146 AEIYNTNSGTTPTDLNPGVSMNTNLVFPDM 176
Db 160 AQIALDNDIDAVMDININPGNTVDVSLVYDMP 190

RESULT 12
0931R6 STAM
ID 0931R6_STAM PRELIMINARY; PRT; 6713 AA.
AC 0931R6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, last annotation update)
DE Hypothetical protein ebnA.
GN Name=ebnA; OrderedLocustNames=SAV1434;
OS Staphylococcus aureus (strain M50 / ATCC 700639).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
```

```
RL Lancet 357:1225-1240(2001).
DR EMBL; BA000017; BAB57596.1; -; Genomic_DNA.
DR SMR; Q931R6; 1-125.
DR InterPro; IPR011439; DUF1542.
DR InterPro; IPR011490; FIVAR.
DR InterPro; IPR002988; GA.
DR InterPro; IPR006530; YD.
DR pfam; PF07564; DUF1542; 8.
DR pfam; PF07554; FIVAR; 44.
DR pfam; PF07554; FIVAR; 44.
DR pfam; PF01468; GA; 46.
DR TIGRfams; TIGR01643; YD repeat 2x; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 6713 AA; 722314 MW; BBC0536AC341BFS CRC64;

Query Match
Best Local Similarity 24.4%; Score 113.5; DB 2; Length 6713;
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;

QY 1 NTNVNIRGNTPRNAST--VSATSLVNSTNSQVAKQEQNSSTSPAKHSTNSLQHQHOA 58
Db 899 NONVYDADSKRNNAVTAQVTAEGILNKQTGNTSKADVNDALNVTRAKKALNGAEKNR 958
QY 59 ATTSSSOSKLRYPHTYGVGDLEIRVNSLQQVSVGVYDG--GETANGAFWVINTIT 115
Db 959 NAKTSATNTINGLPVLTLQQDNLGHQYBQANV--VGYNQVYDKNTLNTMGALRTSI 1016
QY 116 RNDGSTPEVVDGIFHLQNLNGVYQPDSTAEIYNTNSGTTPTDLNPGVSMNTNLVFPDM 175
Db 1017 QNDNTT--KTSQNYLADSDSNKNY---NTA---VNNANGVYINATNRP--NMDAVALNDM 1066

RESULT 13
099U54 STAM
ID 099U54_STAM PRELIMINARY; PRT; 6713 AA.
AC 099U54;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
DE EbnA protein.
GN Name=ebnA; OrderedLocustNames=SA1267;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
```

```
Query Match
Best Local Similarity 24.4%; Score 113.5; DB 2; Length 6713;
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;
```


; Sequence 19725, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,768
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19725
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19725

Query Match 9.7%; Score 104; DB 2; Length 417;
Best Local Similarity 25.4%; Pred. No. 0.045; Indels 54; Gaps 11;
Matches 53; Conservative 32; Mismatches 70;

QY 9 NTPRNASTVSA-----TTSLVNS-----TNSOVAKQONSGST 41
DB 36 NTQRILNNSASALNTSLQRLSTGSRINSKXDAAGLQIANRLTSQVNGLVATKXANGI 95
QY 42 SPA-----AKSTSLQAHQ-----QAATSSQSKRLYIPHTYGVGDLEIRNSLQ 91
DB 96 SLAQTAEGALQOSTIILQRMRLSLQSANSGNSDBERTAL-----NGEVKQLQ---KEIDR 148
QY 92 VSVGVYDGIQGTANGAFWYINITIRNDGSTPMQVVD-GIFHL--ONLNGVYQPDSTARI 148
DB 149 ISNTTFGGRKLLDSFGVASFQV--GSANETIISVIGIDKMSASLNGTTFKADGGAAV 205
QY 149 YANTNSGTFPTDLPNGVS--MTNLVFPDM 175
DB 206 TAATASGTV--DIAIGITGSAVNVKVM 232

RESULT 3
US-10-172-502-18
; Sequence 18, Application US/10172502
; Patent No. 6841154
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-18

Query Match 9.2%; Score 99; DB 2; Length 485;
Best Local Similarity 24.5%; Pred. No. 0.18;
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;
QY 1 NTVVIRKNTPRNASTVSAATSLVNST-----NSSOVAKQONSGSTSPAHKS-----TNS 50
DB 11 NTQSETVGN--QNSTTIEASTADSTSVTKNSSSV--QTSNSDFTVSSSEKSTSTNS 66
QY 51 LQAHQAATSSQSKRLYIPHTYGVGDLEIRNSLQOVKSVGYDGIQGTANGAFWV 110
DB 67 TSNQEKLTSTSESTSSK-----NTTSSSDTKSYA----- 96

QY 111 INTIRNDGSTPMQVVDGIFHLQNLNGVYQPDSTARIYANTNSGTFPTDLPNGVSMTN 170
DB 97 -----STSSTEQPI-----NTSTNGSTAS--NNTSQSTTPSSSVNLNKTSTTS 136
QY 171 LVFDMPTFTYGVGHQHSYLSVSMGFPSGDETTVAL 206
DB 137 TSTAPVKLTSTSESTSSK-----AMSTFASAAITTAIV 164

RESULT 4
US-10-172-502-4
; Sequence 4, Application US/10172502
; Patent No. 6841154
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2283
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-4

Query Match 9.2%; Score 99; DB 2; Length 2283;
Best Local Similarity 24.5%; Pred. No. 1.9;
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVVIRKNTPRNASTVSAATSLVNST-----NSSOVAKQONSGSTSPAHKS-----TNS 50
DB 113 NTQSETVGN--QNSTTIEASTADSTSVTKNSSSV--QTSNSDFTVSSSEKSTSTNS 168
QY 51 LQAHQAATSSQSKRLYIPHTYGVGDLEIRNSLQOVKSVGYDGIQGTANGAFWV 110
DB 169 TSNQEKLTSTSESTSSK-----NTTSSSDTKSYA----- 198
QY 111 INTIRNDGSTPMQVVDGIFHLQNLNGVYQPDSTARIYANTNSGTFPTDLPNGVSMTN 170
DB 199 -----STSSTEQPI-----NTSTNGSTAS--NNTSQSTTPSSSVNLNKTSTTS 238
QY 171 LVFDMPTFTYGVGHQHSYLSVSMGFPSGDETTVAL 206
DB 239 TSTAPVKLTSTSESTSSK-----AMSTFASAAITTAIV 266

RESULT 5
US-09-248-796A-15273
; Sequence 15273, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15273
; LENGTH: 994
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15273

Query Match 9.0%; Score 97; DB 2; Length 994;

TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match
Best Local Similarity 21.3%; Score 94; DB 2; Length 1375;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;

QY 10 TPRNASTVATSLVNSTNSQVAKOEONSTSPAKSTNSLOHQAQAT 60
DB 67 TATDTSTATSATSQATATVTDVSTNGSTNTATATAFVVKPTTSEAKTDSNDKIT 126
QY 61 TSSSQSKL---RYIPFH-----TYKVGDLIRVNSLQOVKSVGYDGIETANGAF 108
DB 127 TSKAVNRLTATGKFPVANNNTAHPKVTVDKIVPIKPIGKLGKQPSLSODDIALGN--- 183
QY 109 WYINT-----IRNGSTPMEVVDGIFHLQNLNGNVYQDPSTAEIYANT---NSGTI 157
DB 184 -VKNIKRYVNGKYYKKEGTLQKN-----VALNTNGKTFPFDETGALSNNTLPKKGNI 236
QY 158 PTDLPNGVSMNTNLVFDMPDFWYTGVGQHY 188
DB 237 TNDNTNSFAQYNOYVS--TDVANFEHV-DHY 265

RESULT 9
US-09-270-767-31859

Sequence 31859, Application US/09270767

Patient No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 31859

LENGTH: 313

TYPE: PRT

ORGANISM: Drosophila melanogaster

US-09-270-767-31859

Query Match
Best Local Similarity 25.7%; Score 93.5; DB 2; Length 313;
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;

QY 5 MIRGTPRNASTVATSL-----VNSTNSQVAKO---EONSTSPAKSTNSLOH 53
DB 142 VVSTVSTSLNTTASTASTAGHRTKDSIHNSSSGAGSQFYCELNSTSRHNSLDRDVH 201
QY 54 AQOHA-ATTSSSQSKLRYIPFTYGVKVDLEIRVNSLQOVKSVGYDGIETANGAFWYIN 112
DB 202 HQHASLMSASSSVSNLSLDSNSGGR-----QSGKLTHSGMWG-NGLSKHNGT----- 249
QY 113 ITIRNDGSTPMEVVDGIFHLQNLNGNVYQDPDS 144
DB 250 -----TNGGSP-----HINAGLVQPS 266

RESULT 10
US-09-270-767-47076

Sequence 47076, Application US/09270767

Patient No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 47076

LENGTH: 313

TYPE: PRT

ORGANISM: Drosophila melanogaster

US-09-270-767-47076

Query Match
Best Local Similarity 25.7%; Score 93.5; DB 2; Length 313;
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;

QY 5 MIRGTPRNASTVATSL-----VNSTNSQVAKO---EONSTSPAKSTNSLOH 53
DB 142 VVSTVSTSLNTTASTASTAGHRTKDSIHNSSSGAGSQFYCELNSTSRHNSLDRDVH 201
QY 54 AQOHA-ATTSSSQSKLRYIPFTYGVKVDLEIRVNSLQOVKSVGYDGIETANGAFWYIN 112
DB 202 HQHASLMSASSSVSNLSLDSNSGGR-----QSGKLTHSGMWG-NGLSKHNGT----- 249
QY 113 ITIRNDGSTPMEVVDGIFHLQNLNGNVYQDPDS 144
DB 250 -----TNGGSP-----HINAGLVQPS 266

RESULT 11
US-09-041-991A-8

Sequence 8, Application US/09041991A

Patient No. 6107278

GENERAL INFORMATION:

APPLICANT: Schnepf, H. Ernest

APPLICANT: Narva, Kenneth E.

APPLICANT: Muller-Cohn, Judy

TITLE OF INVENTION: Toxins Active Against Peets

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: USA

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/041,991A

FILING DATE: 13-MAR-1998

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

TELEPHONE/DOCKET NUMBER: MA-709

TELEPHONE: (352) 375-8100

TELEFAX: (352) 372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 633 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-041-991A-8

Query Match
Best Local Similarity 23.0%; Score 93.5; DB 2; Length 633;
Matches 42; Conservative 33; Mismatches 73; Indels 35; Gaps 8;

QY 1 NTVWIRKN-TPRNASTVATSLVNSTNSQVAKOEONSTSPAKSTNSLOHQAQAA 59
DB 472 NNTAVHENGHTMTHLAPDNTGFTTSPTHATVNNQTTFTSEKGNQSDSLRFSQ---- 527
QY 60 TTSSSQSKLRYIPFTYGVKVDLEIRVNSLQOVKSVGYDGIETANGAFWV---INITIR 116

Db 528 -----SNTARY- TLKNGNSVNLKRS-----SIGNSTIRTINGRYTASNVTTN 576
QY 117 NDGSTMEVVDGIFHLQNLNGNVQPDSTAEIYATNNGTIPDLNPGISMTTNLVEDMP 176
Db 577 NDG-----VNDNGARFSDINIGNVASSNS--DVPDLIN--VTLNSGTQFDLM 620
QY 177 DFM 179
Db 621 NIM 623

RESULT 12
US-09-608-533A-8
Sequence 8, Application US/09608533A
Patent No. 653464
GENERAL INFORMATION:
APPLICANT: Schuepf, H. Ernest
Narva, Kenneth E.
Miller-Cohn, Judy
TITLE OF INVENTION: Toxins Active Against Pests
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/608,533A
FILING DATE: 30-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/041,991
FILING DATE: 13-MARCH-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sanders, Jay M.
REGISTRATION NUMBER: 39,355
REFERENCE/DOCKET NUMBER: WA-709D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 633 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-608-533A-8

Query Match 8.7%; Score 93.5; DB 2; Length 633;
Best Local Similarity 23.0%; Pred. No. 0.98;
Matches 42; Conservative 33; Mismatches 73; Indels 35; Gaps 8;

QY 1 NTVYMRGN-TPRNASVTSATSLVNSTNSGVAKQONSSTSPAHKSTNSIQHQA 59
Db 472 NNIVAVHNGETWHLAPEDNTGFTISPIHATQVNNQTRFPISEKFGNSDSIRPQ---- 527
QY 60 TTSSQSKLRYLPFTHYGVGDLERVNSLQOVKSVGVDGETANGAFW---INITR 116
Db 528 -----SNTARY- TLKNGNSVNLKRS-----SIGNSTIRTINGRYTASNVTTN 576
QY 117 NDGSTMEVVDGIFHLQNLNGNVQPDSTAEIYATNNGTIPDLNPGISMTTNLVEDMP 176
Db 577 NDG-----VNDNGARFSDINIGNVASSNS--DVPDLIN--VTLNSGTQFDLM 620

QY 177 DFM 179
Db 621 NIM 623

RESULT 13
US-09-538-092-1254
Sequence 1254, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Glot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuraPatSeqFormatter Version 0.9
SEQ ID NO 1254
LENGTH: 784
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (0) --(0)
OTHER INFORMATION: Polypeptide Accession Number Q02446
US-09-538-092-1254

Query Match 8.7%; Score 93.5; DB 2; Length 784;
Best Local Similarity 34.0%; Pred. No. 1.4;
Matches 34; Conservative 17; Mismatches 32; Indels 17; Gaps 4;

QY 9 NTPRNASV-----SATS LVNSTNSGVAKQONSSTSPAHKSTNSIQHQA 56
Db 306 NTVYMRGN-TPRNASVTSATSLVNSTNSGVAKQONSSTSPAHKSTNSIQHQA 56
QY 57 QAAATSSQSKLRYLPFTHYGVGDLERVNSLQOVKSVG 96
Db 365 PATESBAQSSQ--LHANG-MQNDQONSLSLQOVYLG 400

RESULT 14
US-09-949-016-10170
Sequence 10170, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10170
LENGTH: 786
TYPE: PRT
ORGANISM: Human
US-09-949-016-10170

Query Match 8.7%; Score 93.5; DB 2; Length 786;
Best Local Similarity 33.0%; Pred. No. 1.4;

Matches 33; Conservative 17; Mismatches 33; Indels 17; Gaps 3;

QY NTPRNASTV-----SATSIVNSTSSQVAKOEONSSSTPAHKSTNSLQHAOH 56

Db 308 NTTTSASTMPSPSSSTCTTASTSTLTSSTLVSSADTQGYASTS-ASSERTIESSQT 366

QY 57 QAAATSSSQSKLRYPFTYTGKVDLEIRVNSLQOVKSVG 96

Db 367 PATTESEAQSSSQLOP---NGMOMAQOQNSLQOVQIVG 402

RESULT 15

US-09-949-016-6805

; Sequence 6805, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6805

; LENGTH: 1228

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-6805

Query Match 8.6%; Score 92.5; DB 2; Length 1228;

Best Local Similarity 22.0%; Pred. No. 3.3;

Matches 47; Conservative 27; Mismatches 79; Indels 61; Gaps 9;

QY 16 TVSATTSIVNST--NSQVAKOEONSSSTPAHKSTN----- 49

Db 1018 TVNLTTLVLIGRTQRTMDNIYPEEYSSCS-RHPCONGGTCINGRTSFTCACRHPPTGDNC 1076

QY 50 SLOHAHQATSSSQSKLRYPF-----HTYKVGDLERVNSLQOVKSVDIGET 103

Db 1077 TIKLYEENLAPDFSRGSTRAPMVAFPASHTYGTIPGPIIFNNL---DVYGASTYP 1132

QY 104 ANGAF-----WVINITIRNDGS--TPMEVVDGI-----FHLQNLNGVY----- 140

Db 1133 RTGKRIRIPYLGYYVRYKTTIESFAHISGLVVDGIDKLAFAESRNINSRHCORVLTGDAL 1192

QY 141 -OPDSTAEIYANTNSGTIPTDLNPGVSMTNLVF 173

Db 1193 LEINYQGEVWMLAKGITIPAKFPVYTFESGYLLY 1226

Search completed: April 7, 2006, 16:25:20
Job time : 31.4813 secs

RESULTS 2
US-09-815-242-5912

```
/ Sequence 5912, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5912
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-5912

Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

Qy      74 HTG---KVGDLIRVNSLQOVKSYGDIGETANGAFWVINITRNDGSTPMEVVDGIF 130
      81 HRIGETVKGDLLEVTVNSVETKWSGSLAPTNAGIFVADVITIKNGKCALTTIDSSMF 140
      131 HLQNLNGVYOPDSTAEIYAN-TNSGTIFPD-----LNPVSMNTNLVFPDMPD 177
      141 KLS-GDKTFRADNTGSMNSQSDNGSIENSFPLQRIINPDSTAQKIVFDVSE 192
Db

RESULT 3
US-09-815-242-12872
/ Sequence 12872, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13146
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-13146

Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
```

```
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12872
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-12872

Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

Qy      74 HTG---KVGDLIRVNSLQOVKSYGDIGETANGAFWVINITRNDGSTPMEVVDGIF 130
      81 HRIGETVKGDLLEVTVNSVETKWSGSLAPTNAGIFVADVITIKNGKCALTTIDSSMF 140
      131 HLQNLNGVYOPDSTAEIYAN-TNSGTIFPD-----LNPVSMNTNLVFPDMPD 177
      141 KLS-GDKTFRADNTGSMNSQSDNGSIENSFPLQRIINPDSTAQKIVFDVSE 192
Db

RESULT 4
US-09-815-242-13146
/ Sequence 13146, Application US/09815242
/ Patent No. US2002061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ FILE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13146
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-13146

Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
```

```
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
Qy 74 HTYG--KVGDEIRVNSIQOVKSVGYDGGTANGAFVYVITTRNDGSTMEVVDGIF 130
Db 81 HKIGETVKNXGDLLEVYVNSVETKSVGPSLAPTNAGKIFVAVVITTKNKGKELTTIDSSWF 140
Qy 131 HLOINGNVYOPDSTAETIYAN-TNSGTIPTD-----LNPVSMTNVLVFDMPD 177
Db 141 KLS-KGDKTFEADNTGMSANQSDNGSIENSFFLQRINDPDSFAQCKIVFVDSR 192

RESULT 5
US-10-282-122A-44046
; Sequence 44046, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykied, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 44046
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-44046

Query Match 12.6%; Score 135; DB 4; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
Qy 74 HTYG--KVGDEIRVNSIQOVKSVGYDGGTANGAFVYVITTRNDGSTMEVVDGIF 130
Db 81 HKIGETVKNXGDLLEVYVNSVETKSVGPSLAPTNAGKIFVAVVITTKNKGKELTTIDSSWF 140
Qy 131 HLOINGNVYOPDSTAETIYAN-TNSGTIPTD-----LNPVSMTNVLVFDMPD 177
Db 141 KLS-KGDKTFEADNTGMSANQSDNGSIENSFFLQRINDPDSFAQCKIVFVDSR 192
```

```
RESULT 6
US-10-282-122A-43811
; Sequence 43811, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykied, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 43811
; LENGTH: 6713
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43811

Query Match 10.6%; Score 113.5; DB 4; Length 6713;
Best Local Similarity 24.4%; Pred. No. 3.6;
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;
Qy 1 NTVMIRGNTPRNST--VSATSLVNSTSSQVAKQONSSTSPAHKSTNSLOHQAHOA 58
Db 899 NONVYLDADSKRNATVQAVTAABGILNKQGTGSKADVDNALNAVTRAKALNAGABNLR 958
Qy 59 ATTSSOSKAKYIPPHYTGXGDLIRVNSIQOVKSVGYDGI---GETRANGAFVYVITTI 115
Db 959 NAKTSATNTINGLPVLTQLOKDNLKHVEQAQNV--VGVNGVKDKGNTLNTAMGALRTSI 1016
Qy 116 RNDGSTMEVVDGIFHLQNLGNVYOPDSTAETIYANTSGTIPDNLNPVSMTNVLVFDMPD 175
Db 1017 QNDVTT--KTSQNTYLDASDSKKNY---NTA---VNNANGVYINMTNNE--NMDANAINDM 1066

RESULT 7
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5639
LENGTH: 2086
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5639

Query Match 10.0%; Score 107.5; DB 3; Length 2086;
Best Local Similarity 24.2%; Pred. No. 2.6;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

1 NTVMIRGNTPRNAST--VSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLOHQAQOA 58
1323 NONTYDADESKRMAYTQAVTAABGLNKQCGANTSKADVNDALNTVTAKAALNGAENLR 1382
59 ATTSSQSKLRYIPFHTYGVKGDLEIRVNSLQOVKSV-GYDGIETGANGAFVINITRN 117
1383 NTKTSATNTINGLPULTQLOKDNLKHQVEQAQNVAGVGDKXGNTLNTAKALATSTION 1442
118 DGSFMEVVDGIFHLQNLNGNVYQPDSTAEIYANTNSGTFPTDLNPGVSMTTNLVFD 175
1443 DNTT--KTSQNYLADSDSNKNY---NTA---VNNANGVINVTNMP--NMDANALNGM 1490

US-09-815-242-12610
Sequence 12610, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12610
LENGTH: 5795
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12610

Query Match 10.0%; Score 107.5; DB 3; Length 5795;
Best Local Similarity 24.2%; Pred. No. 11;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

1 NTVMIRGNTPRNAST--VSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLOHQAQOA 58
3600 NONTYDADESKRMAYTQAVTAABGLNKQCGANTSKADVNDALNTVTAKAALNGAENLR 3659
59 ATTSSQSKLRYIPFHTYGVKGDLEIRVNSLQOVKSV-GYDGIETGANGAFVINITRN 117
3660 NTKTSATNTINGLPULTQLOKDNLKHQVEQAQNVAGVGDKXGNTLNTAKALATSTION 3719
118 DGSFMEVVDGIFHLQNLNGNVYQPDSTAEIYANTNSGTFPTDLNPGVSMTTNLVFD 175
3720 DNTT--KTSQNYLADSDSNKNY---NTA---VNNANGVINVTNMP--NMDANALNGM 3767

US-10-437-963-131231
Sequence 131231, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boudharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 131231
LENGTH: 1020
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_33316C.1.dep
US-10-437-963-131231

US-10-437-963-131231
Sequence 131231, Application US/10437963
Publication No. US20040123343A1

Query Match 9.8%; Score 105.5; DB 4; Length 1020;
Best Local Similarity 25.1%; Pred. No. 1.5;
Matches 51; Conservative 32; Mismatches 63; Indels 57; Gaps 11;

1 NTVMIRGNTPRNASTVSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLOHQAQOA 60
573 NTFYSSGK-----SASVLSMSTLHSSQGS-----NSNMQIGQ----- 607
61 TSSQSKLRYIPFHTYGVKGDLEIRVNSLQOVKSV-----VGYDGIETG--TANGAFW 109

Db 608 ---GQLKQ-----YNAAGALSLVSNKSTQYEAANEQKRYKYYIGQDNGSDPDQAAGAYI 658

Qy 110 VI--NITLRNDGSPMEYVDGIF----HLQMLNGVYQPDST--AEIYANTN--SGTIP 158

Db 659 FRPNQTPFPIKIDQVPLIVLRSGILDEYH--QQINPMYIQLINRYKKGADYVEIEFIVGP 717

Qy 159 TDLNPGVSMNTNLVDFMDPEFTY 181

Db 718 VDDGNKEKELSTEVYTNMAFNKTF 740

```

RESULT 10
US-10-156-761-7751
; Sequence 7751, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIORITY FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7751
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7751

```

Query Match	9.8%	Score 105	DB 4	Length 2386
Best Local Similarity	22.2%	Pred. No. 5.3		
Matches	52	Conservative	30	Mismatches 92
				Indels 60
				Gaps 10
QY	14	ASTVSATTSLVNSTNSQVAKQKONSTSPAKHSTNSLQHAQHAATTSSSQSLRYLP	73	
Db	1364	SKVEHETASAYTTAYGYTAGGELAQITDPNGNNTLYYDMAQRKTTDDPDAGLSSEY	1423	
QY	74	HTYKVGLELRNWSLQQYKSVGVGIDGIELANG-----AFWVINTIRDSSTPEMV	126	
Db	1424	NENQGVSGTTATTNDVQVLTLYYGDNLSRAVSRSAGDELAWWDDPAATGKGK-QIT	1481	
QY	127	DGIFHLQNLNGVYQPOSTAEIYANTNSG-----TIPTDLNPQVS--MTTNL	171	
Db	1482	SAV--SRDASNTY-----TTKKGKFKDBGRPLNTVTLPITYN-GLADYITTSV	1528	
QY	172	VFDNPDFA-----TYGHVGHRTSLVANSQFPGS---DETTY	204	
Db	1529	TYDAADHTTSVSPAGKLAERKVTYYDDVGQFTRLITSSIG--GTAYIDWTTY	1580	

RESULT 11
US-09-815-242-12713
Sequence 12713, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in

```

, TITLE OF INVENTION: Prokaryotes
, FILE REFERENCE: ELITRA.011A
, CURRENT APPLICATION NUMBER: US/09/815,242
, CURRENT FILING DATE: 2001-03-21
, PRIOR APPLICATION NUMBER: 60/191,078
, PRIOR FILING DATE: 2000-03-21
, PRIOR APPLICATION NUMBER: 60/206,848
, PRIOR FILING DATE: 2000-05-23
, PRIOR APPLICATION NUMBER: 60/207,727
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: 60/242,578
, PRIOR FILING DATE: 2000-10-23
, PRIOR APPLICATION NUMBER: 60/253,625
, PRIOR FILING DATE: 2000-11-27
, PRIOR APPLICATION NUMBER: 60/257,931
, PRIOR FILING DATE: 2000-12-22
, PRIOR APPLICATION NUMBER: 60/269,308
, PRIOR FILING DATE: 2001-02-16
, NUMBER OF SEQ ID NOS: 14110
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 12713
, LENGTH: 2344
, TYPE: prt
, ORGANISM: Staphylococcus aureus
US-09-815-242-12713

```

Query Match	9.64;	Score 103;	DB 3;	Length 2344;
Best Local Similarity	25.58;	Pred. No. 8;		
Matches	55;	Conservative 18;	Ident. 71;	Gaps 9.
Qy	1	NTVMYIGKNTPERNAATVSATTSILVNST-----NSSQVAKOEQNSSTSPAHKS-----TNS	50	
Db	101	NTGSEIVGN--QNSTIDASTSTADSTSVTKNSSSV--QTSNDSPTVSEKSEENVTSSTNS	156	
Qy	51	LQHQHQAAATTSSSQSKLRYPFTHTYKGVGLIRANSLOQVSVGVDIGETANGARW	110	
Db	157	TSNQOEKLTSTSTSEKTSK-----NTTSSSDPTKSV-----	185	
Qy	111	INTIRNDSTPMEVNVGIFHLQNLGNVYQPDSTAIYANTNSGIIPTDLNPGVSMTN	170	
Db	186	-----TSTSTSEQPI-----NTSNQSTAS--NNTSGSTPTSLANKTSTTS	226	
Qy	171	LVFDMPPDFTYGHVGHSLVASKGFGSGDETTAL	206	
Db	227	TSTAPVRLRTFSRL-----ANSTASATTTAL	254	

```

RESULT 12
US-10-125-692-14
Sequence 14, Application US/10125692
Publication No. US2003004429A1
GENERAL INFORMATION:
APPLICANT: Aderem, Alan
APPLICANT: Hayaishi, Punataka
APPLICANT: Smith, Kelly D.
APPLICANT: Underhill, David M.
APPLICANT: Ozimek, Adrian
TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
TITLE OF INVENTION: of Use
FILE REFERENCE: P-Is 5155
CURRENT APPLICATION NUMBER: US/10/125,692
CURRENT FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: US 60/285,477
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 394
TYPE: PRT
ORGANISM: P. aeruginosa
US-10-125-692-14

```

Query Match 9.3%; Score 100; DB 4; Length 394;

Best Local Similarity 24.4%; Pred. No. 1.3;
Matches 51; Conservative 32; Mismatches 72; Indels 54; Gaps 10;

QY NTPRRASTVSA-----TTSLVNS-----TNSQVAKQONST 41
DB 13 NTRNLNNSASLNTSLORLSGSRINSKDDAQLANRLTSOVNGLAVATKANOGI 72
QY 42 SPA-----HKSTNSLOHACH---QAAATSSOSKLRYPHTYGVGLERVNSIQ 91
DB 73 SLAQTBAGALQOSTTILQRMRLSLQANGSNDSEBRTALN-----GEAKQLOKELDR 125
QY 92 VKSVGYDGISETANGAFVYVITITINDGSTPEVVD-GIFHL--ONLNGNVYQDPSTAEI 148
DB 126 ISNTTFGGRKLLDGSFGVASFGV--GSAANEIISVIGIDEMASBSLNGTFKADGGGAV 182
QY 149 YANTNSGTITPDLPNGVS--MTTNLVFDM 175
DB 183 TAAIASGTV--DIAIGITGSAVNKVKDM 209

RESULT 13
US-10-991-347-14

; Sequence 14, Application US/10991347
; Publication No. US20050147627A1
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan
; APPLICANT: Hayashi, Fumitaka
; APPLICANT: Smith, Kelly D.
; APPLICANT: Underhill, David M.
; APPLICANT: Ozinsky, Adrian
; TITLE OF INVENTION: Toll-Like Receptor 5 ligands and methods
; FILE REFERENCE: 66661-122
; CURRENT APPLICATION NUMBER: US/10/991,347
; CURRENT FILING DATE: 2004-11-16
; PRIOR APPLICATION NUMBER: US 60/285,477
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 10/125,692
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PaedSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 394
; TYPE: PRT
; ORGANISM: P. aeruginosa
US-10-991-347-14

Query Match 9.3%; Score 100; DB 5; Length 394;
Best Local Similarity 24.4%; Pred. No. 1.3;
Matches 51; Conservative 32; Mismatches 72; Indels 54; Gaps 10;

QY NTPRRASTVSA-----TTSLVNS-----TNSQVAKQONST 41
DB 13 NTRNLNNSASLNTSLORLSGSRINSKDDAQLANRLTSOVNGLAVATKANOGI 72
QY 42 SPA-----HKSTNSLOHACH---QAAATSSOSKLRYPHTYGVGLERVNSIQ 91
DB 73 SLAQTBAGALQOSTTILQRMRLSLQANGSNDSEBRTALN-----GEAKQLOKELDR 125
QY 92 VKSVGYDGISETANGAFVYVITITINDGSTPEVVD-GIFHL--ONLNGNVYQDPSTAEI 148
DB 126 ISNTTFGGRKLLDGSFGVASFGV--GSAANEIISVIGIDEMASBSLNGTFKADGGGAV 182
QY 149 YANTNSGTITPDLPNGVS--MTTNLVFDM 175
DB 183 TAAIASGTV--DIAIGITGSAVNKVKDM 209

RESULT 14
US-10-172-502-18
; Sequence 18, Application US/10172502
; Publication No. US20030185833A1
; GENERAL INFORMATION:

; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-172-502-18

Query Match 9.2%; Score 99; DB 4; Length 485;
Best Local Similarity 24.5%; Pred. No. 2.1;
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVVMIRGNTPRNASTVSATTSLVNST-----NSSQVAKQONSTSPARKS-----TNS 50
DB 11 NTQSETVGN--QNSTTIASTADSTVTNKS SV--QTSNSDVTSSSEKSKVTSTTNS 66
QY 51 LQHAQHQAATTSSOSKLRYPHTYGVGLERVNSLQOVKSVGYDGISETANGAFV 110
DB 67 TSNQGEKLTSTSESTSSK-----NTSSSDPKSYA----- 96
QY 111 INTRNDGSTPEVVDGIFHLQNLNGNVYQDPSTAEIYANTNSGTITPDLPNGVSMTTN 170
DB 97 -----STSTEQPI-----NTSTNOSTAS--NNTSOSTTSSVNLNKTSTTS 136
QY 171 LVFDMPEFTYGHVQHSYLSVSMGFSGDEFTVAL 206
DB 137 TSTAPVKLRITSRL-----AMSTFASATTTAV 164

RESULT 15
US-11-020-509-18

; Sequence 18, Application US/11020509
; Publication No. US20050106648A1
; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .
; FILE REFERENCE: P07263US02/BAS
; CURRENT APPLICATION NUMBER: US/11/020,509
; CURRENT FILING DATE: 2004-12-27
; PRIOR APPLICATION NUMBER: US 10/172,502
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-11-020-509-18

Query Match 9.2%; Score 99; DB 6; Length 485;
Best Local Similarity 24.5%; Pred. No. 2.1;
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVVMIRGNTPRNASTVSATTSLVNST-----NSSQVAKQONSTSPARKS-----TNS 50
DB 11 NTQSETVGN--QNSTTIASTADSTVTNKS SV--QTSNSDVTSSSEKSKVTSTTNS 66
QY 51 LQHAQHQAATTSSOSKLRYPHTYGVGLERVNSLQOVKSVGYDGISETANGAFV 110
DB 67 TSNQGEKLTSTSESTSSK-----NTSSSDPKSYA----- 96
QY 111 INTRNDGSTPEVVDGIFHLQNLNGNVYQDPSTAEIYANTNSGTITPDLPNGVSMTTN 170
DB 97 -----STSTEQPI-----NTSTNOSTAS--NNTSOSTTSSVNLNKTSTTS 136

Oy 171 LVFDMDFWYGHVGOHYSLVASMGPFSGDETTVAL 206
Db 137 TSTAPVKLRTFSRL-----AMSTFASAAITTVAV 164

Search completed: April 7, 2006, 16:44:39
Job time : 100.998 secs

.

.

This Page Blank (uspto)

Db 378 ASSSVASSSSSSSSSSAAA 396

RESULT 2

US-10-517-939-74
Sequence 74, Application US/10517939
Publication No. US20060003433A1
GENERAL INFORMATION:
APPLICANT: Steer, Brian
APPLICANT: Callen, Walter
APPLICANT: Healey, Shaun
APPLICANT: Hazlewood, Geoff
APPLICANT: Wu, Di
APPLICANT: Blum, David
APPLICANT: Besehlahlan, Alireza
TITLE OF INVENTION: XYNAMASES, NUCLEIC ACIDS ENCODING THEM
TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
FILE REFERENCE: 564462007901
CURRENT APPLICATION NUMBER: US/10/517,939
CURRENT FILING DATE: 2004-12-13
PRIOR APPLICATION NUMBER: PCT/US03/19153
PRIOR FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: 60/389,299
NUMBER OF SEQ ID NOS: 380
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 74
LENGTH: 1484
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Obtained from an environmental sample
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)... (21)
US-10-517-939-74

Query Match
Best Local Similarity 23.7%; Pred. No. 2.2;
Matches 47; Conservative 28; Mismatches 67; Indels 56; Gaps 10;

QY 17 VSATSLVNSTNSQVAKQONSSTPAKSTNSLQHAQOAATSSOSKLRYPFH 75
DB 661 ISAWYKLSIPATSOQLSTOVNGGTA---SYNNL-----QKISTEDGVNKLGGTYR 711
QY 76 YKRVGD--LEIRVNSLQOVKSVGYDGIGETANGAFVNIITIRNDGSTMEVVDGIFHLQ 133
DB 712 YSSVDEFLIYVES-----SNNSYASFYIDITFEYSSTGSGPIEVED---LT 755
QY 134 NUNGNYVQPD-----STAEIYANT-----NSGTIPTDLNPGVSMITNLVFDMPD- 177
DB 756 PIK-DYIQDDFLIGNAVASDLEGNRLKLLNHHNVYTAENAMKPDQAINAKQDPDIDE 814
QY 178 -----FMTYGHV 184
DB 815 NALVDKVLDPQGLQLGHV 832

RESULT 3

US-11-096-568A-22863
Sequence 22863, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Theydy
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 22863
LENGTH: 295

TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(295)
OTHER INFORMATION: Ceres Seq. ID no. 12410415
US-11-096-568A-22863

Query Match
Best Local Similarity 19.0%; Pred. No. 0.51; Length 295;
Matches 36; Conservative 30; Mismatches 72; Indels 51; Gaps 7;

QY 15 STVSATSLVNSTNSQVAKQONSSTPAKSTNSLQHAQOAATSSOSKLRYPFH 74
DB 130 SYMMDTATVANTANAVAPSPMDTLAQF-----SAGHYGAAATNNCSSM----- 177
QY 75 TYGKVGDLIRVNSLQOVKSVGYDGIGETANGAFVNIITIRNDGST-PMEVVDGIFHLQ 133
DB 178 -----ESPST-WPTSEAVEQENWVPLRAMPDPAQVY 209
QY 134 NUNGNYVQPDSTAEIYANTNSGTIPTDLNPGVSMITNLVFDMPD--PDFMTYGHVGOHSLV 191
DB 210 NFLSIFDPDTSGLH--QMLKAMPVDIETALLMRNLNMLRSPDF-----EQHRL 261
QY 192 ASMGFFGSD 200
DB 262 SSYSTYGAD 270

RESULT 4

US-11-096-568A-22862
Sequence 22862, Application US/11096568A
Publication No. US20060048240A1
GENERAL INFORMATION:
APPLICANT: Alexandrov, Nikolai et al.
TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides
TITLE OF INVENTION: Theydy
FILE REFERENCE: 2750-1592PUS2
CURRENT APPLICATION NUMBER: US/11/096,568A
CURRENT FILING DATE: 2005-04-01
NUMBER OF SEQ ID NOS: 34471
SEQ ID NO 22862
LENGTH: 333
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(333)
OTHER INFORMATION: Ceres Seq. ID no. 12410414
US-11-096-568A-22862

Query Match
Best Local Similarity 19.0%; Pred. No. 0.6; Length 333;
Matches 36; Conservative 30; Mismatches 72; Indels 51; Gaps 7;

QY 15 STVSATSLVNSTNSQVAKQONSSTPAKSTNSLQHAQOAATSSOSKLRYPFH 74
DB 168 SYMMDTATVANTANAVAPSPMDTLAQF-----SAGHYGAAATNNCSSM----- 215
QY 75 TYGKVGDLIRVNSLQOVKSVGYDGIGETANGAFVNIITIRNDGST-PMEVVDGIFHLQ 133
DB 216 -----ESPST-WPTSEAVEQENWVPLRAMPDPAQVY 247
QY 134 NUNGNYVQPDSTAEIYANTNSGTIPTDLNPGVSMITNLVFDMPD--PDFMTYGHVGOHSLV 191
DB 248 NFLSIFDPDTSGLH--QMLKAMPVDIETALLMRNLNMLRSPDF-----EQHRL 299
QY 192 ASMGFFGSD 200
DB 300 SSYSTYGAD 308

RESULT 5

US-11-096-568A-22861
; Sequence 22861, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; PRIORITY FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 22861
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(336)
; OTHER INFORMATION: Ceres Seq. ID no. 12410413
US-11-096-568A-22861

Query Match 8.4%; Score 90.5; DB 7; Length 336;
Best Local Similarity 19.0%; Pred. No. 0.61;
Matches 36; Conservative 30; Mismatches 72; Indels 51; Gaps 7;

QY 15 STVSATSLVNSTSSQVAKQKQNSSTSPAKSTNSLQHQHQATSSQSKRTYIPPH 74
DB 171 SVPMWDTAVRNTNANVAVPSWMDTLAQPF-----SAGHVQGAALATNNCSSM----- 218
QY 75 TVGKYVPLEIRVNSLQGVKSVGVDIGETANGAFWVINTITINDST-PMEVVDGIFHLQ 133
DB 219 -----ESPSTG-WPTSAVQKQNVPPRLAMPDPAQY 250
QY 134 NINGNVYQDSTAEIYVANTNSGTTPTDLPVGSMTNLVFDW--PDFTYGVHGQHYSLV 191
DB 251 NFIAGSIPDDTSGHL--QWLKAMDPIVDIETALLIKRNLNLSMNLRSPPF-----EQHRRLL 302
QY 192 ASMGFRGSD 200
DB 303 SSSYGGAD 311

RESULT 6
US-11-130-821-4
; Sequence 4, Application US/11330821
; Publication No. US20060019275A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: Cowan, David
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved in Sensory
; FILE REFERENCE: 02307E-084210US
; CURRENT APPLICATION NUMBER: US/11/130,821
; PRIORITY FILING DATE: 2005-05-16
; PRIOR APPLICATION NUMBER: US/09/361,630
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/094,464
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse taste cell polypeptide (TCP) #2 amino acid
US-11-130-821-4

Query Match 8.3%; Score 89.5; DB 7; Length 729;

Best Local Similarity 28.5%; Pred. No. 2.2;
Matches 37; Conservative 14; Mismatches 44; Indels 35; Gaps 6;

QY 1 NTVMIRGNTPRNASTVSATTS-----TVSATSLVNSTN-SSQVAKQKQNS-----TSPAKSTNS 50
DB 243 NTVALDTPGVSRGADSDPQTPTSTTDSFKTSNLPQALQPSHGMLMFTSPIHPTLS 302
QY 51 LQH-----AQHQATSSS-----QSKRTYIPHTYK---VGDLEIR 85
DB 303 LQHFSSPSTASTSSGFTSSVHADPTLASTLPHPQDMSLQDLSPTSGRSHHTHSVTR 362
QY 86 VNSLQGVSV 95
DB 363 INSNRFTYAV 372

RESULT 7
US-11-052-554A-216
; Sequence 216, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIORITY FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patent In version 3.3
; SEQ ID NO 216
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Streptococcus mutans UA159
US-11-052-554A-216

Query Match 8.1%; Score 87.5; DB 7; Length 507;
Best Local Similarity 21.1%; Pred. No. 2.1;
Matches 56; Conservative 38; Mismatches 93; Indels 79; Gaps 14;

QY 2 TVTMRGNTPRNASTVSATTS--LVNSTSSQVAKQKQNSSTSPAKST-----NSLQH 53
DB 48 TVSQADGNDPQSTSVQOETAPOCTKTSQSSDATVDSBSATSPDEQTVSGQDSSSSQ 107
QY 54 AQHQATSSSQSKRTYIPHTYGVGDLEIRVNSLQ----- 90
DB 108 IDQTIADNRSDSD--HISKTSATTEDEKRVNSAKQTAATNNOPTRYSAKDAYGNS 165
QY 91 -----QVKSYGVDIGETANGAFWVINTITINDOSTPM-----EVV-----DG 128
DB 166 NFNKLTLPGRKANVADVTTNGVRD-----EIVVNDPSAPYVRNAEIAKLEY 216
QY 129 IFHLQNLGNVYQF-----DSTAEIYVANTNSGTTPTDLPVGSMTNLVFDWDFMTY--- 181
DB 217 LTELKLNIN-NIAIPVPSVDQWQKXAQDRAHEANEKQ-GLDHDTNL--PIPNNLTWVAE 272
QY 182 -GHVGQHYSLVASM--GP-FGSDETT 203
DB 273 DGHLDMSISQSKSQEGYTLASDKAT 298

RESULT 8
US-11-052-554A-95
; Sequence 95, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A

;; CURRENT APPLICATION NUMBER: US/11/052,554A
;; CURRENT FILING DATE: 2005-02-07
;; PRIOR APPLICATION NUMBER: US 60/589,227
;; PRIOR FILING DATE: 2004-07-20
;; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
;; PRIOR FILING DATE: 2004-02-06
;; NUMBER OF SEQ ID NOS: 763
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 95
;; LENGTH: 1237
;; TYPE: PRT
;; ORGANISM: Helicobacter pylori J99
US-11-052-554A-95

Query Match 8.0%; Score 85.5; DB 7; Length 1237;
Best Local Similarity 21.4%; Pred. No. 11;
Matches 48; Conservative 33; Mismatches 58; Indels 85; Gaps 11;

QY 17 VSATSLVNSTSSQVAKOE-----ONSSTSPAKSTNSLQ---HAHQ-----A 58
DB 482 VSQVSVYNSLKTSENIQNAKILCNNGSGSTSPCNSSGGLISGNAQLQNLSPTN 541
QY 59 ATTSSQ-----SKRYIPFHYGVGDLIRVNSLQOVKSVGYD-----GIGETAN 105
DB 542 GTTNTQAKSNASKTKKAM-----VWVNBEEKAKTTNFQSSGPTQSSNSTVM 589
QY 106 GAFWVINTIRNDGSTPEV-----VDGIFHLQNLNGVYQPDST----- 145
DB 590 GA-----LNTVQLQVNSFPQSGISAPQNGENNTQAMNALTNSNPNGNSQNLITNNODL 646
QY 146 -----AEIY--ANTNSGTIPDNL-----PGVSMITTN 170
DB 647 RIQLAFNYQLINTINQVPTMNLINQSOOTQOTSGSASTTN 690

RESULT 9
US-11-087-099-3488
;; Sequence 3488, Application US/11087099
;; Publication No. US20060041961A1
;; GENERAL INFORMATION:
;; APPLICANT: Abad, Mark S. et al.
;; TITLE OF INVENTION: Genes and Uses for Plant Improvement
;; FILE REFERENCE: 38-21(53450)B EP
;; CURRENT APPLICATION NUMBER: US/11/087,099
;; CURRENT FILING DATE: 2005-03-22
;; NUMBER OF SEQ ID NOS: 12464
;; SEQ ID NO 3488
;; LENGTH: 500
;; TYPE: PRT
;; ORGANISM: Clostridium acetobutylicum
US-11-087-099-3488

Query Match 7.9%; Score 85; DB 7; Length 500;
Best Local Similarity 25.1%; Pred. No. 3.5;
Matches 42; Conservative 22; Mismatches 65; Indels 38; Gaps 5;
QY 10 TPRNST-----VSATSLVNSTSSQVAKQONSGSTSPAKSTNSLQHAQQAATSS 64
DB 199 TFSNASKNKINTSSDINSIVSDNSGKITATAEGSGTATA-----TTTAAS 245
QY 65 QSKLRYPFHYGVGDLIRVNSLQOVKSVGYDGETANGAFWVINTIRNDGSTPE 124
DB 246 DSQKTATCVVTVNKRTPITFKDNL-----EIS-----IRMNINKPTGSLKE 288
QY 125 VVDGIFHLQNLNGVYQPD-----STABIYANTNSGTIPDNLNGVSMT 168
DB 289 DVQGITSFYNTNPNHILDGIELNLBLEFTPSNTPIDLPKLSLT 335

RESULT 10
US-10-055-877-140
;; Sequence 140, Application US/10055877
;; Publication No. US20050288241A1

;; GENERAL INFORMATION:
;; APPLICANT: Decristofaro, Marc
;; APPLICANT: Padigar, Muralidhara
;; APPLICANT: Miller, Charles
;; APPLICANT: Tchernev, Velizar
;; APPLICANT: Zhong, Mei
;; APPLICANT: Anderson, David
;; APPLICANT: Ballinger, Robert
;; APPLICANT: Gerlach, Valerie
;; APPLICANT: Spytek, Kimberly
;; APPLICANT: Kaceli, Luca
;; APPLICANT: Kerkuda, Rameesh
;; APPLICANT: Guo, Xiaojia
;; APPLICANT: Zernhusen, Bryan
;; APPLICANT: Andrew, David
;; APPLICANT: Mezas, Peter
;; APPLICANT: Paturajan, Meera
;; APPLICANT: Burgess, Catherine
;; APPLICANT: Eileen, Andrew
;; APPLICANT: Wolenc, Adam
;; APPLICANT: Baumgartner, Jason
;; APPLICANT: Shimkets, Richard
;; APPLICANT: Gusev, Vladimir
;; APPLICANT: Vermet, Corine
;; APPLICANT: Taupier Jr., Raymond
;; APPLICANT: Pena, Carol
;; APPLICANT: Shenoy, Sureeh
;; APPLICANT: Li, Li
;; APPLICANT: Casman, Stacie
;; APPLICANT: Boldog, Ference
;; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
;; FILE REFERENCE: 21402-251
;; CURRENT APPLICATION NUMBER: US/10/055,877
;; CURRENT FILING DATE: 2002-01-22
;; PRIOR APPLICATION NUMBER: 60/262,892
;; PRIOR FILING DATE: 2001-01-19
;; PRIOR APPLICATION NUMBER: 60/263,598
;; PRIOR FILING DATE: 2001-01-23
;; PRIOR APPLICATION NUMBER: 60/263,799
;; PRIOR FILING DATE: 2001-01-24
;; PRIOR APPLICATION NUMBER: 60/264,117
;; PRIOR FILING DATE: 2001-01-25
;; PRIOR APPLICATION NUMBER: 60/264,139
;; PRIOR FILING DATE: 2001-01-25
;; PRIOR APPLICATION NUMBER: 60/264,478
;; PRIOR FILING DATE: 2001-01-26
;; PRIOR APPLICATION NUMBER: 60/263,351
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: 60/272,870
;; PRIOR FILING DATE: 2001-03-02
;; PRIOR APPLICATION NUMBER: 60/275,990
;; PRIOR FILING DATE: 2001-03-14
;; PRIOR APPLICATION NUMBER: 60/275,927
;; PRIOR FILING DATE: 2001-03-14
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 512
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 140
;; LENGTH: 1765
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-10-055-877-140

Query Match 7.9%; Score 85; DB 6; Length 1765;
Best Local Similarity 24.6%; Pred. No. 20;
Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;
QY 18 SATSLVNSTSSQVAKQONSGSTSPAKSTNSLQHAQQAATSSQSKLRYPFHYG 77
DB 670 NANNVLSKSKSLTFLHLLAODRVAEVLVNGCAHV-----DAOTKGTTPHAYGC 721
QY 78 KVGDLEIRVNSLQOVKSV-----GYDGIGETA-NGAFWVINTIRNDGSTPEHVDGI 129

Db 722 HYGNIKI-VNFILOHSAKVNAKTKNGYTAHQAAQOQHTTHIIIVLQNNAS-PNEL----- 775
QY 130 FHLQNNNGN-----VYQPS-----TAIYANTNSGTTPTDLPNGVSWTTLVDP 174
Db 776 ----TVNGNTALAIARRLGIVSVDTLKVYTEIMTT--TTEKHKNVPEITNBEVLD 828
QY 175 MPD 177
Db 829 MSD 831

RESULT 11
US-10-055-877-141
; Sequence 141, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Murajidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Ratelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zernhusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Paturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eissen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Verneet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 141

; LENGTH: 1940
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-055-877-141

Query Match
7.9%; Score 85; DB 6; Length 1940;
Best local Similarity 24.6%; Pred. No. 23;
Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;

QY 18 SATSLVNSTSSQVAKQEQNSSTSPAHKSTNSLOHAQOAMTTSSQSKRLRYIPHTY 77
Db 670 NANNVLSNKSGLTPLHLAAQEDRVNVAEVLVNGAHV-----DAQRTMGYTPILVGC 721
QY 78 KYGDLKIRVNSLQVKS-----GYDQIGETA-NGARVIVNITIRNDGSTMEYVDGI 129
Db 722 HYGNIKI-VNFILOHSAKVNAKTKNGYTAHQAAQOQHTTHIIIVLQNNAS-PNEL----- 775
QY 130 FHLQNNNGN-----VYQPS-----TAIYANTNSGTTPTDLPNGVSWTTLVDP 174
Db 776 ----TVNGNTALAIARRLGIVSVDTLKVYTEIMTT--TTEKHKNVPEITNBEVLD 828
QY 175 MPD 177
Db 829 MSD 831

RESULT 12
US-11-096-568A-29105
; Sequence 29105, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; PRIOR FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 29105
; LENGTH: 946
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE: Arabidopsis thaliana
; NAME/KEY: misc.feature
; LOCATION: (1)..(946)
; OTHER INFORMATION: Ceres Seq. ID no. 4263811
US-11-096-568A-29105

Query Match
7.9%; Score 84.5; DB 7; Length 946;
Best local Similarity 21.7%; Pred. No. 9.3;
Matches 43; Conservative 32; Mismatches 82; Indels 41; Gaps 7;

QY 11 PNNASTVSATSLVNSTSSQVAKQEQNSSTSPAHKSTNSLOHAQOAMTTSSQSKRLRY 70
Db 59 PNNSSVILSLSNFPLNSNSSFPLVYCNLTLESILVSNRRLSIEBGFYTCNERLIALKH 118
QY 71 IPRHT-----YGVGDLKIRVNSLQ-QVYSGVDGIGETANGAFVYINITI-RN 117
Db 119 INFSTNKSTSPRGFGFSLAVLDPSSHVLSGVNDYDGFGLVQLRS-----LNTSPRL 173
QY 118 DGSITMEYVDGIFHLQNNNGNNGVYQDSTAETIYANTNSGTTPTDLPNGVSWTTLVDPMD 177
Db 174 TGSVEVHLTKSLDKL-----EVSNDLSGTTIPGIDYDGLT--LIDSLD 216
QY 178 FWTYGVGHVSLVSMG 195
Db 217 NQANG-----SIPSLG 228

RESULT 13
US-11-096-568A-29104
; Sequence 29104, Application US/11096568A
; Publication No. US20060048240A1

